Population genetic structure of a planktonic diatom in the Gulf of naples : *Pseudo-nitzschia multistriata*

Thesis

The Open University

POPULATION GENETIC STRUCTURE OF A PLANKTONIC DIATOM IN THE GULF OF NAPLES: *PSEUDO-NITZSCHIA MULTISTRIATA*

SYLVIE VAÏANA MARIE TESSON

MASTER DEGREE IN OCEANOGRAPHY AND MARINE ENVIRONMENTS, UNIVERSITY PIERRE AND MARIE CURIE - PARIS VI

DOCTOR OF PHILOSOPHY

OPEN UNIVERSITY OF LONDON
STAZIONE ZOOLOGICA ANTON DOHRN,
NAPLES, ITALY

OCTOBER 2010

Director of Studies: Dr Wiebe H.C.F. Kooistra

Laboratory of Ecology and Evolution of Plankton

Stazione Zoologica Anton Dohrn, Naples, Italy

Internal Supervisor: Dr Gabriele Procaccini

Laboratory of Functional and Evolutionary Ecology

Stazione Zoologica Anton Dohrn, Naples, Italy

External Supervisor: Prof. Paul K. Hayes

University of Portsmouth, United Kingdom

« Quand les mouettes ont pieds, il est temps de virer »

(Almanach du Marin Breton)

When the seagull are walking, it is time to change the direction;

Meaning that when there is an impasse, a solution must exist.

TABLE OF CONTENTS

ABSTRACT

The planktonic diatom *Pseudo-nitzschia multistriata* appeared in plankton samples taken at the MareChiara station in the Gulf of Naples in 1995 and has appeared ever since in autumn, and since 2004 also in summer. This thesis focuses on population genetic structure of the species using the Internal Transcribed Spacers in the rDNA as well as seven polymorphic microsatellites.

Cells isolated from plankton samples taken in 2008 and 2009 were grown into strains from which ITS-sequences and microsatellite fingerprints were gathered. Microsatellite screening uncovered two populations in sympatry in 2008, only one of which reappeared in 2009. The ITS sequences of the strains were of distinct types, called A-type, B-type and a mixed type (A/B). The two populations also showed distinct ITS-type proportions, supporting their distinctness. Intra-population proportions of ITS-types among strains were similar in summer and autumn 2009, indicating stability over the growth season. But proportions differed between 2008 and 2009, suggesting substantial immigration the second year. These results suggest that, *P. multistriata* in the Mediterranean consists of a dynamic patchwork of populations.

Microsatellite inheritance was studied over the vegetative and sexual parts of the biennial life cycle of *P. multistriata*. Substitutions were recorded in strains re-screened after several months of maintenance. Longer microsatellites changed more frequently than shorter ones and di-nucleotide ones more than tri-nucleotide ones. Strains interbred independently from population assignment. Comparisons of parental and offspring genotypes showed microsatellite inheritance patterns generally following Mendelian rules. Genetic identity of sister cells emerging from zygotes formed on the same parental gametangia indicate that one of the two nuclei resulting from the first meiotic division is eliminated. Each maternal gamete fuses with a paternal gamete, resulting in two zygotes, which, in this case, were genotypically identical. Thus, mutations boost allele diversity and biennial sexual reproduction boosts ITS-genotype variability and genotypic diversity.

ACKNOWLEDGEMENTS

My acknowledgements go to the people that turned the Ph.D.-project into a reality.

I am grateful to my director of studies, Dr *Wiebe HCF Kooistra*, for trusting and supporting me during the thesis. For his instructive teaching, his particular views on analysing, writing and presenting science. To Dr *Gabriele Procaccini*, my internal supervisor, for his help, support and good advice.

To the *Open University* of London for the three-years studentship; To the *Stazione Zoologica Anton Dohrn* of Naples and *Ecology and Evolution of Plankton* laboratory for having been my hosts and given me financial and technical support to fulfill this project.

To people that contributed to this project: Crew of the Vettoria and Management and Ecology of Coastal Areas for collecting weekly net samples. Dr *Diana Sarno* for providing LTER-MC data on *Pseudo-nitzschia multistriata* counts. SBM-molecular biology service, Drs *Elio Biffali* and *Marco Borra* for their advice in molecular biology and their availability; to *Elvira*, *Lello*, *Raimondo* and *Fabrizio* for assistance with samples, their hospitality and the resting eleven o'clock pizza. *Carmen Minucci* for her precious biomolecular support. I will not forget the hours spent together elaborating the massive collection of DNA material, nor the labyrinth of furniture orders. Dr *Christophe Brunet* for his help in clarifying obscure side of statistic approaches and *Gandhi Forlani* for his technical support with culture chambers. I am grateful to Dr *Marina Montresor* for her precious advices and succulent ethnical lunches and to Prof *Paul K. Hayes*, my external supervisor, for his helpful comments during the thesis elaboration and writing.

Dr *Sonia Quijano-Scheggia* from Institut de Ciències del Mar (Barcelone, Spain) and Dr *Ana Amorim* from Instituto de Oceanografia (University of Lisbon, Portugal) for supplying culture materials of *Pseudo-nitzschia multistriata.*

To Marine Genomics Europe (Alfred Wagner Institute (DE), Station Biologique de Roscoff (FR)), GenEco (University of Lund (SE)), SItEMICRO (CNR-ISE, Verbania (IT)), OU-SZN (Stazione Zoologica Anton Dohrn (IT)), and workshops: Eurocean (Stazione Zoologica Anton Dohrn (IT)), Marplan (Marbef) & AQUAPARADOX (Observatoire Oceanologique de Villefranche-sur-mer (FR)), that enabled me to learn genetic analysis approaches and techniques, encountering peers and enlarging my interest in genetic field.

I am grateful to Dr *Domenico D'Alelio* for our great scientific discussion on *P. multistriata*. To *Emanuella, Chiara, Ornella, Daniella,* and *M.Valeria* for our biomolecular encounters, helpful discussions, and the good moments spent together. To *Christophe Legrand* for his collaboration on the fascinating *P. multistriata* sex story in revealing genetic inheritance. To *Melania Salzano* and *Donatella Forio*, thesis-students, for their pertinent questions and their interest in biology. Thank Donatella to continue the adventure with 2010 samples!

I would like to further thanks my lab mates, *Giuseppe, Eleonora, Deepak, Raffaele, M.Valeria, Maria-Letizia, Vasco, Lucia, Anna I., Isabella, Gianluca* and *Gauri* for their friendship, our interesting open discussion, resting tea time and soothing lunch time. And also *Ferdinando, Agostina, Benedetto, Rosario, Andrea, Cecilia, Simona, Alessandra, Ilaria I., Antonio, Luisa, Leopoldo, Claudia, Elena, Frauke, Nicole, Ankita, George, Francesca* and *Imma*. Thanks guys for sharing good and bad moments; our Aperitivi at Belledone, barbecues, Christmas and pizza parties were precious. To *Gaya, Agostina, Cecilia* and *Dorris*, my flat mates, for supporting me and sharing Naples life and handmade cakes.

A huge thank goes to my family, especially to François & Anne-Marie, Stéphane & Audrey, GP, for their lifetime of love and support, their judgment strength, philosophy, perspicacity and sense of humour. A particular thought goes to GM that has passed away a few months ago.

I am glad to have had the opportunity to live and discover Naples; an alive theatre city, rich in culture and culinary tradition. Thanks to its tasty pizza, mozzarella and diverse pasta.

LIST OF ABBREVIATIONS

© / ®, ™	Copyright / Registered, unregistered TradeMark
°C	Degree Celcius
BET	Ethidium bromide
BLAST	Basic Local Alignment Search Tool
bp	Base pair
ca.	*circa* (approximately, about)
CTAB	Hexadecyltrimethylammonium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphate
e.g.	*Exempli gratia* (for example)
EDTA	Ethylene Diamine Tetra Acetic acid
EEP	Laboratory of Ecology and Evolution of Plankton
et al.	*et alia* (and other)
etc.	*et cetera* (and so on)
EtOH	Ethanol
FA	Fragment analyser
Fw	Forward primer
GoG	Gulf of Gaeta
GoN	Gulf of Naples
GoS	Gulf of Salerno
i.e.	*id est* (that is to say)
ITS	Internal Transcribed Spacer region of the rDNA
L:D	Light:Dark hour photoperiod
LM	Light Microscope
LSU	Large Subunit Ribosomal DNA
LTER-MC	Long Term Ecological Research station – MareChiara
N, S, E, W	North, Sud, East, West (cardinal point)
PCR	Polymerase Chain Reaction
Ph.D.	Doctor of Philosophy
PNm	Microsatellite locus designed for *P.multistriata*
rDNA	Ribosomal DNA
Rev	Reverse primer
SEVAG	Chloroform Isoamyl alcohol 24:1
SZN	Stazione Zoologica Anton Dohrn
T°a	Temperature of annealing
T°m	Temperature of melting
TBE	Tris Borate EDTA
WGA	Whole Genome Amplification

SI units

μg	Microgram (10^{-6} kilogram, SI unit)
μl & ml	Microlitre & Millilitre (10^{-6} & 10^{-3}litre, SI unit)
μm & nm	Micrometer & Nanometer (10^{-6} & 10^{-9} meter, SI unit)
μmol photon.m^{-2}.s^{-1}	Unit of irradiance
min	Minute
pmol	Picomoles (10^{-12} mole, SI unit)
psu	Practical salinity unit (1 psu = 1 g of NaCl per kg of seawater)
rpm	Revolution per minute (unit of frequency of rotation)
u.μl^{-1}	Unit per microlitre

LIST OF FIGURES

LIST OF TABLES

LIST OF APPENDICES

CHAPTER ONE

GENERAL INTRODUCTION

1.1. Introduction

This thesis explores the genetic structure of populations of the pennate diatom *Pseudo-nitzschia multistriata* at the Long Term Ecological Research MareChiara (LTER-MC) sample site in the Gulf of Naples. To place the work into context, each of the different facets of the thesis are introduced in this Chapter.

First, information is presented about the physical oceanographic properties of the Gulf of Naples (GoN), its seasonality and patterns of current flow. Then characteristics of the phytoplankton in the GoN will be presented. Since the study species, *P. multistriata,* is a diatom, a general introduction to the diatoms will be provided, including their diversity, evolutionary history, life cycle and their ecological importance in the plankton. This is followed by an introduction to the genus *Pseudo-nitzschia*, its characteristics, taxonomic diversity and societal relevance. From this body of knowledge, the research questions have been defined, which lead to the hypotheses tested in this study: the hypotheses and questions emerged from previous work carried out in the laboratory of Ecology and Evolution of Plankton (EEP) at the Stazione Zoologica Anton Dohrn (SZN). With the questions defined, the procedures and methodologies used to address them will be presented and it will be explained why these particular ones have been chosen.

1.2. The Gulf of Naples.

1.2.1. Geography

The Gulf of Naples (GoN) is located on the western side of the Italian peninsula, and opens to the Tyrrhenian Sea (Western Mediterranean) on the South-eastern side (Fig. 1.1). It is enclosed on the North-western side by the peninsula of Miseno, and the Islands of Procida and Ischia, on the North-Eastern side by Naples, its hinterland and Vesuvius, and on the South-eastern side by the Peninsula of Sorrento and the Island of Capri. The Gulf is deep with only a narrow shallow coastal zone. The bottom drops rapidly to oceanic depths immediately beyond Ischia and Capri (Fig. 1.1). The Gulf of Gaeta (GoG) is located North-west of the GoN and the Gulf of Salerno (GoS) to the South-East. Both the Gulf of Gaeta and the Gulf of Salerno are more exposed and open to the Tyrrhenian Sea than the GoN.

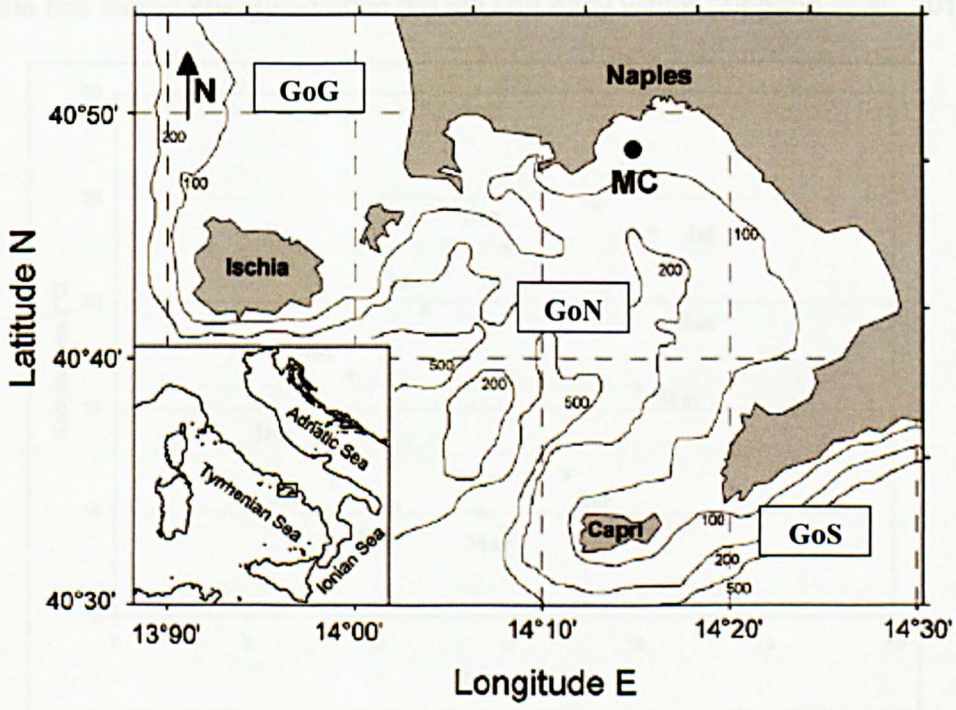


Figure 1.1: Map of the Gulf of Naples showing the location of the LTER MareChiara sampling station (40°48.5'N, 14°15'E; (Zingone et al., 2010)). Contours show the depth in meters. Gulfs of Gaeta (GoG), Naples (GoN) and Salerno (GoS).

1.2.2. Temperature and day-length

Figure 1.2 shows a plot of the average monthly surface water temperature, measured from 1984 to 2009, against the day-length over the annual cycle at Long Term Ecological Research station MareChiara (LTER-MC). The winter and summer solstices, that is, the dates of minimum and maximum amount of energy absorbed by the sea surface waters from incoming radiation, are respectively at the end of December (daylength ~ 9.5 hours) and at the end of June (~ 14.5 hours). The minimum and maximum water temperatures, recorded respectively in February-March (14 °C ± 1) and August (26 °C ± 1.5), lag two months behind the solstices because the sheer mass of the seawater functions as a buffer, requiring a huge amount of energy, and thus time, to warm up in spring and early summer and to dissipate this stored energy again in the fall and early winter (Zingone *et al.*, 2010).

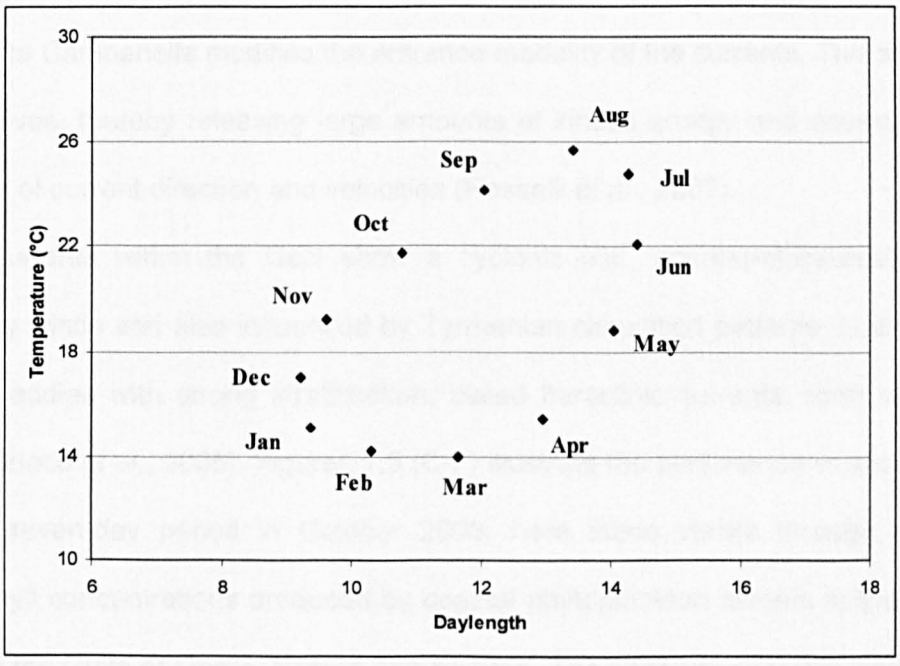


Figure 1.2: Average day-length plotted against monthly surface water temperatures (0-10m) at the LTER-MC (EEP-data).

1.2.3. Currents, winds and nutrients

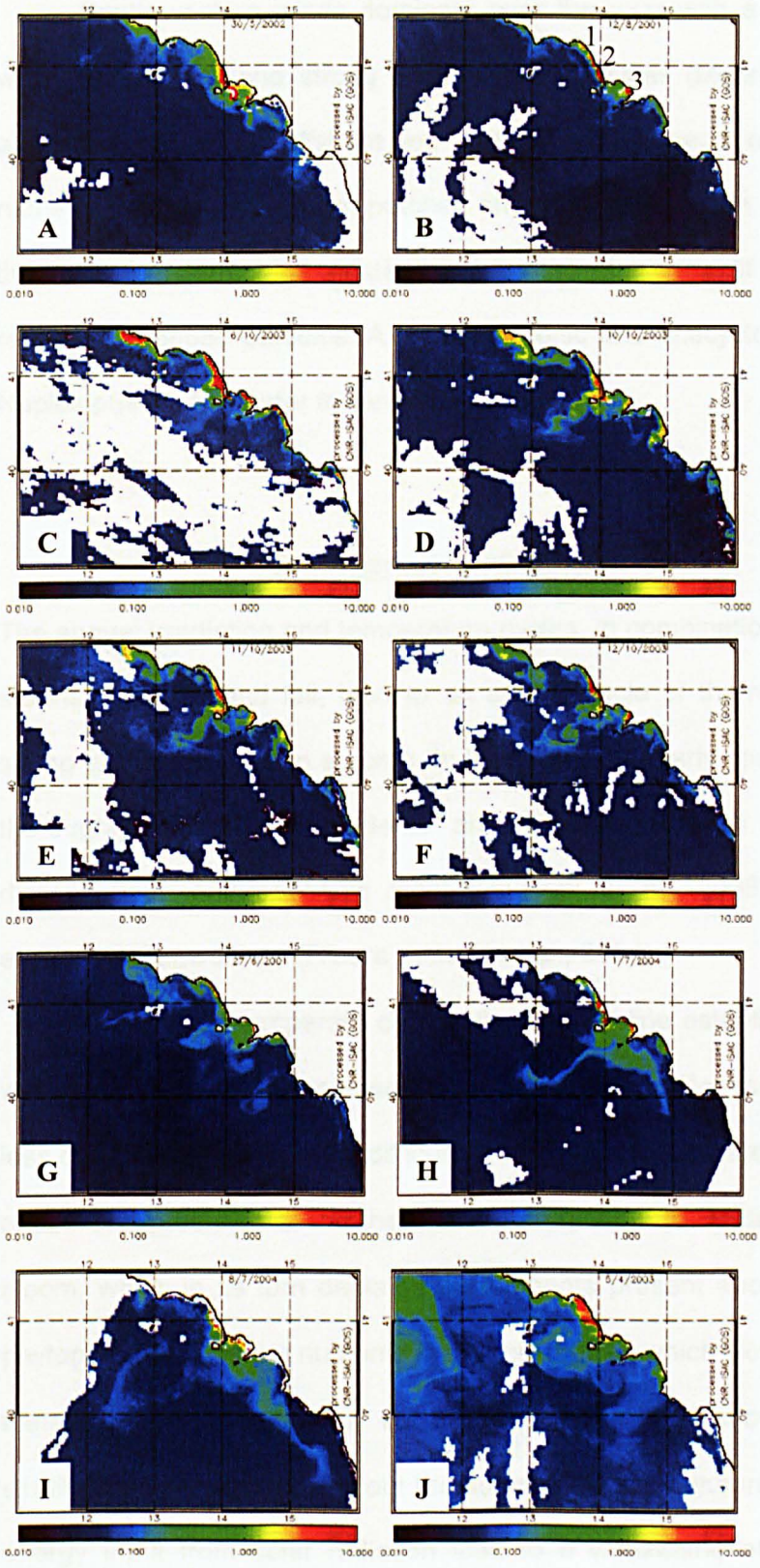
In comparison to the currents in the open oceans, currents in the central Mediterranean are relatively weak because of the small tidal amplitudes and the semi-enclosed nature of the basin. Strong currents occasionally occur locally, (*e.g.* in narrow straits and around promontories) and are generated by atmospheric pressure differences and wind stress. Water movement close to the shore is due to a complex coastal topography and to wind stress, which depends on wind direction and strength.

The GoN represents a semi-enclosed system with its own circulation, wind stress and bathymetry whereas the Gulf of Gaeta (GoG) and the Gulf of Salerno (GoS) are more open to the Tyrrhenian Sea (Fig. 1.1). Over the year in the GoN, currents run along the NW-SE axis, except in summer, where smaller currents and episodes of rapid reversal N-S wards current occur in presence of atmospheric perturbation (Gravili *et al.*, 2001). A sill at a depth of 80 m in the strait between Capri and Punta Campanella modifies the entrance modality of the currents. This sill blocks deep waves, thereby releasing large amounts of kinetic energy and causing abrupt variation of current direction and velocities (Rosselli *et al.*, 2007).

Currents within the GoN show a cyclonic (*i.e.*, counter-clockwise) pattern driven by winds and also influenced by Tyrrhenian circulation patterns. Local rapidly rotating eddies with strong stratification, called baroclinic currents, form along the coast (Grieco *et al.*, 2005). Figures 1.3 (C-F) illustrate the persistence of such eddies over a seven-day period in October 2003, here made visible through elevated chlorophyll concentrations produced by coastal phytoplankton blooms in the surface water of the Gulfs of Gaeta, Naples and Salerno. The changes over this single week provide an insight into the current velocities and patterns often observed in the fall.

Note that a dense bloom on 06/10/03 at the mouth of the Volturno River has moved at 10/10/03 to Procida and Ischia and has become part of an eddy system South-West of Ischia, which can be traced over the next two days. The south-western tip of eddies is pulled to the north-west by the north-west-wards offshore current. The chlorophyll patterns suggest that eddies are a local phenomenon (Fig. 1.3 A) and that they do not move over long distances along the coastline (Fig. 1.3 G). Yet, exceptions occur. Figures 1.3 (H-I) illustrate a plume of coastal water from the GoN spreading south-east-wards along the Italian coast over a few hundred kilometres in eight days.

Figure 1.3: Satellite images showing the distribution of chlorophyll in part of the eastern Tyrrhenian Sea, the Gulf of Gaeta (GoG), the Gulf of Naples (GoN) and the Gulf of Salerno (GoS) along the South-western coast of the Italian peninsula (for orientation and scale compare with Fig. 1.1). The colour scale indicates the chlorophyll concentration estimated from spatial detection (0.010 to 10.000 $mg \cdot m^{-3}$). White indicates land; white spots correspond to missing data due to the presence of clouds. (Satellite images, CNR-ISAC) Fig. 1.3 A: A water mass rich in chlorophyll (red) from the GoG enters the GoN and water mass (green) moves from the GoN through the channel between Capri and the Peninsula of Sorrento into the GoS (30/05/2002). Fig. 1.3 B: Localised phytoplankton blooms in river plumes as indicated by high chlorophyll concentrations (12/08/2001); numbers indicate the mouths of the rivers Garigliano (1), Volturno (2), Sarno (3). Fig. 1.3 C-F: Chlorophyll patterns observed over a period of seven days in autumn (6 to 12/10/2003) illustrating the persistence of eddies. Fig. 1.3 G-J. Chlorophyll patterns showing small cyclonic eddies mixing coastal water rich in chlorophyll and offshore water poor in chlorophyll (G: 27/06/2001)); eddies drift by Mediterranean currents (H-I: 1-8/07/2004); a large anticyclonic pattern of currents in the GoG and GoS, a cyclonic current pattern in the GoN and eddies influence from outside the immediate coastal zone studied (J: 05/04/2003).


A
B
C
D
E
F
G
H
I
J
1
2
3
processed by CNR-ISAC (GOS)
0.010
0.100
1.000
10.000

South-western winds dominate over the year with a predominance of SSW winds in summer and strong NNE winds in winter (Menna *et al.*, 2007). Winter easterlies induce an offshore jet, coastal currents and offshore gyres that trap nutrients and plankton, form patches and prevent nutrients and plankton release in the GoN. Westerly winds instead induce the disruption of offshore gyres and the release of trapped particles. A mix of cyclonic and anticyclonic gyres in the Gulf of Naples pushes the water towards the coast.

1.2.4. Thermal Stratification

The annual irradiation and temperature cycles, in combination with the occurrence of storms in spring and fall, lead to an annual cycle of thermocline establishment in spring and breakdown in autumn. In the winter and early spring, the water column is thoroughly mixed and nutrients are available through meteorological factors, discontinuous anthropogenic input (Pugnetti *et al.*, 2006), currents, thermocline erosion and upwellings (Ribera d'Alcalà et al., 2004).

In Northern temperate climes, the thermocline establishes as the upper layer warms up through the increased energy input from solar radiation. Warmer water is less dense and therefore be difficult to mix through the water column of *ca.* 14 °C at depth. The presence of the thermocline (from April to September) leads to a spring bloom, which in its turn depletes the nutrients present above the thermocline. The phytoplankton fixes the nutrients in organic matter, which slowly sinks out through the thermocline, and temporarily out of reach to the biota above the thermocline. This stratification persists throughout the summer. In the autumn, storms and diminishing energy input from solar radiation lead to a weakening of the thermocline and a progressive disruption of the stratification till December (Zingone et al., 2010).

Nutrients are again available to permit a shorter autumnal plankton bloom, as long as the diminishing light is sufficient to sustain phytoplankton growth.

Salinity does not change markedly across the annual cycle, ranging between *ca.* 37.5 psu in May and 38.0 psu (± 0.1 psu) in September (Ribera d'Alcalà et al., 2004). The lower values in spring result from a net influx of freshwater due to rain. These differences are not strong enough to affect the temperature driven density stratification.

1.3. Phytoplankton patterns and seasonality in the Gulf of Naples

1.3.1. Phytoplankton blooms

In general, phytoplankton blooms in the temperate zone occur in spring and autumn. They are governed by nutrient availability and external physical factors such as temperature, daily amount of irradiance (depending on day-length, the angle of the sun above the horizon) and the amount of particles in the water column that absorb or scatter the light .

During winter the water column is thoroughly mixed and nutrients are in ample supply. Yet, the phytoplankton fails to bloom because the amount of light that it receives is insufficient to sustain growth. This light limitation is linked to deep mixing of the water column in which the phytoplankton resides and the low irradiation levels penetrating the surface because of short days and the low angle of the sun above the horizon. A spring bloom is made possible because of increasing amounts of light entering the water column and because of a reduction of the mixing due to the establishment of a thermocline. This bloom collapses when nutrients are depleted in the water column above the thermocline, but is also influenced by the increase of

pathogens and grazers, feeding on phytoplankton, in the environment. In the fall, the thermocline commences to disintegrate, but the amount of light is still sufficient for algae to grow, giving rise to an autumnal bloom. Figure 1.4 shows that environmental conditions during spring and autumn differ markedly, as the water temperature in the fall is much higher than in spring, which, in part, accounts for the different species composition of those two blooms.

Phytoplankton distribution is influenced by turbulence, currents (Fig. 1.3) and is heterogeneous and patchy (Grieco et al., 2005). Apart from these physical settings, plankton composition in a bloom depends also on biological interactions among species, ranging from differential grazing (Ribera d'Alcalà et al., 2004) or pathogen attack, allelopathic interactions (Cembella, 2003; Tillmann *et al.*, 2007) and competition for resources (Litchman *et al.*, 2010). Also species-specific factors play a role such as life cycles, the ability to form resting stages, endogenous rhythms and the capacity to adapt the physiology to environmental conditions (Lavaud, 2007; Sgrosso *et al.*, 2001; Suzuki & Johnson, 2001).

1.3.2. Phytoplankton in coastal and offshore waters in the GoN

The Gulf of Naples can be divided into an oligotrophic offshore zone, consisting of Tyrrhenian water, and a eutropic coastal zone affected by run-off from the land (Pugnetti et al., 2006). The oligotrophic offshore water is characterised by very low chlorophyll concentrations in the surface water (a deep blue colour in most of the pictures in Fig. 1.3). Chlorophyll concentrations rise principally during the establishment of a thermocline in spring and during the breakdown of this thermocline in autumn, indicative of a spring bloom and an autumn bloom. Eddies may also influence the chlorophyll distribution and drive it in and/or out the GoN (Fig. 1.3 G-J).

Close to the coast, eutrophic conditions persist year round, resulting in lush phytoplankton growth as illustrated in Fig. 1.3 by the high chlorophyll coastal concentrations throughout the year. Values remain high for two reasons. First, the warm surface water reaches all the way to the shallow bottom resulting in thorough mixing of the entire water column. Hence, organic matter and the nutrients are recycled *in situ* and remain available. Second, the coastal water receives nutrient input through discharge from anthropogenic sources such as the wastewater purification system near Cuma in the Gulf of Gaeta, discharge from the rivers Volturno, Gargliano and Sarno (Fig. 1.3 B) and through incidental runoff during rainstorms (Pugnetti et al., 2006). Late autumnal storms can transport the coastal surface water straight out of the GoN and replace it with oligotrophic Tyrrhenian water mixed in with coastal water from the GoN or from elsewhere.

1.3.3. Phytoplankton recording at the LTER-MC in the GoN

The SZN maintains a long-term ecological research (LTER) station, called MareChiara (MC), situated two nautical miles offshore in the transition zone between the coastal and open Tyrrhenian waters. The bottom is located at 80 m. Environmental data and plankton samples have been collected at this station every 14 days from 1984 until the early summer of 1991. Data collection is now on a weekly basis and recommenced during the winter of 1995. Plankton net samples are gathered by means of horizontal and vertical tows and water samples are taken at a series of depths using a Rosette sampler. Phytoplanktonic diversity and abundance are determinated by light microscopy (LM). Most diatoms and dinoflagellates are identified down to the species level in as far as this is possible using light microscopy.

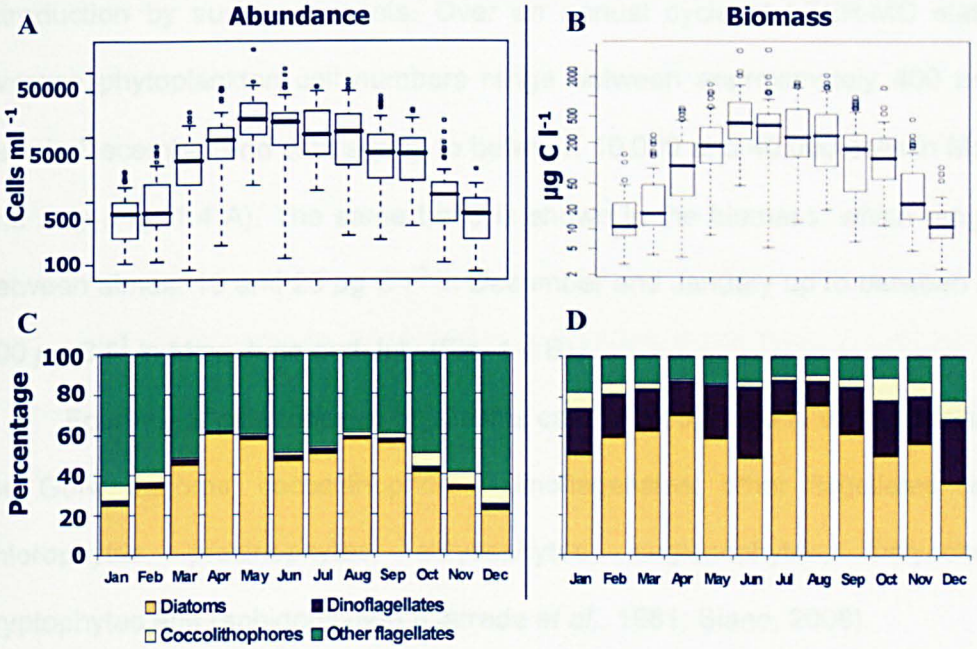


Figure 1.4: Box plots and bar charts showing abundance and biomass of four categories of phytoplankton organisms, namely diatoms, coccolithophores, dinoflagellates and other flagellates (see text). Values are given for each month of the year and are averages measured over the samples taken in each of these months from 1984-1991 (fortnightly samples) to 1995-2006 (weekly samples) in the surface water layer at the LTER-MC station. Fig. 1.4 A: Phytoplankton cell numbers per month. The bold lines represent medians, boxes denote the range of the middle 50% of the values. The horizontal lines indicate the upper and lower 95% limits and the small circles represent upper outliers. Fig. 1.4 B: Phytoplankton biomass per month. Legend as in Fig. 1.4 A. Fig. 1.4 C: Percentage of the cell numbers for each of the four categories of phytoplankton organisms. Fig. 1.4 D: Percentage of the biomass for each of the four categories of phytoplankton organisms.

The pattern of bi-modal distribution of phytoplankton species was observed in the Gulf of Naples at LTER-MC, taking into the account the five most abundant species (Ribera d'Alcalà et al., 2004). Blooms occur as a major peak in spring (about May) and a minor peak in later summer- autumn (about August). However the LTER-MC phytoplankton distribution during the summer is smoother thanks to species introduction by surface currents. Over an annual cycle at LTER-MC station, the average phytoplankton cell numbers range between approximately 400 and 1000 cells in December and January up to between 10,000 and 40,000 cells in May, June and July (Fig. 1.4 A). The same trend is shown in the biomass, which ranges from between almost 10 and 20 μg $C \cdot l^{-1}$ in December and January up to between 200 and 800 μg $C \cdot l^{-1}$ in May, June and July (Fig. 1.4 B).

Four major categories of organisms can be recognized in the phytoplankton of the GoN: diatoms; coccolithophores; dinoflagellates; other flagellates (including chlorophytes, prasinophytes, chrysophytes, euglenophytes, dictyochophytes, cryptophytes and raphidophytes) (Carrada *et al.*, 1981; Siano, 2008).

Numerically the diatoms and the "other" flagellates dominate all the year round whereas the dinoflagellates and coccolithophores constitute only a small percentage of the total cell numbers. Diatoms show a higher percentage in spring and summer; the "other" flagellates dominating from November to February; coccolithophores show their maximum abundance from October and February, (about 4 %) and dinoflagellates account for just 2 % of the counts (Fig. 1.4 C). The use of biomass, to partition the community results in a radically different picture (Fig. 1.4 D): diatoms are still important in the plankton year-round (40 to 80 % of the biomass), with dinoflagellates contributing 10 to 40 %, the "other" flagellates between 10 and 20 %, and coccolithophorids making a very minor contribution. These contrasting patterns

are a result of factors such as cell size differences between the groups and differences in the cell wall. For example, photoautotrophic dinoflagellates are generally large celled organisms whereas "other" flagellates constitute a collection of rather small organisms. Dinoflagellates possess organic cell walls, which count in the total biomass, whereas diatoms have silica cell walls, which do not count in the biomass. Diatoms can also be large, too, but in this group the cells have a large vacuole that does not contribute significantly to the biomass. A major part of the dry weight of coccolithophores does not add to the biomass either because their cell wall consists of calcium carbonate coccoliths.

Long-term trends are also visible in the phytoplankton counts at the LTER-MC. Since 1995 the community of phytoplankton shows a decrease in cell size and an increase in abundances of diatoms and phytoflagellates (Ribera d'Alcalà et al., 2004).

1.4. Diatoms

1.4.1. General structure of the diatom and its cell wall

Pseudo-nitzschia multistriata, the subject species of this study, is a pennate diatom. Diatoms are unicellular protistan autotrophs. The hallmark of the diatom cell is its rigid, compound silica cell wall, called a frustule, which completely encases the protoplasma but that is structured to allow cell growth and division. The frustule consists of two Petri dish shaped valves, and their accompanying sets of bands or rings, called cingular bands or girdle bands (Fig. 1.5).

One set of a valve and cingular bands, called hypotheca, fits perfectly inside the confines of the other set, called epitheca. This cell wall architecture results in a

peculiarity of the diatoms, namely that in the course of the vegetative phase in the diatom life cycle, the average cell size diminishes. Upon each cell division, the old epitheca obtains a new hypotheca, resulting in a daughter cell of the same size as the mother cell. However, the old hypotheca now becomes the epitheca of the other daughter cell, which obtains a new hypotheca inside it. Thus, the daughter cell inheriting the hypotheca of the mother cell will be minutely narrower than the mother cell (Fig. 1.6). As a result, the average size of cells in a cohort decreases and the variance around the average increases with time. This reduction in size during asexual reproduction in diatoms is called the MacDonald-Pfitzer rule (MacDonald, 1869; Pfitzer, 1869). When a certain threshold size is reached the cells become sexually mature. Diatoms exhibit a diplont life cycle with a long vegetative diploid phase (*ca.* two years (D'Alelio *et al.*, 2010)), and a brief phase of sexual reproduction in which initial cell size is re-established. The diatom life cycle will be explained in further detail below (section 1.4.3).

The frustule elements comprise a compound material composed of an organic matrix on which ortho-silicic acid ($SiOH_4$; ca 96.5%) and small amounts of other molecules (ca 1.5% Al_2O_3 or Fe_2O_3) precipitate in an amorphous form (Lee, 1999; Rogall, 1939). The construction of new frustule elements is carried out under weakly acidic conditions in specialized vesicles, called silica deposition vesicles. Various long chain proteins and enzymes that control and shepherd silica precipitation are laid down into a matrix, the shape and local density of which determines silica density and the shape of micro- and nanostructures in the elements (Hildebrand, 2008; Kröger & Poulsen, 2008; Wenzl *et al.*, 2008). When construction is complete, the vesicle membrane fuses with the plasmalemma, thus exuding the element.

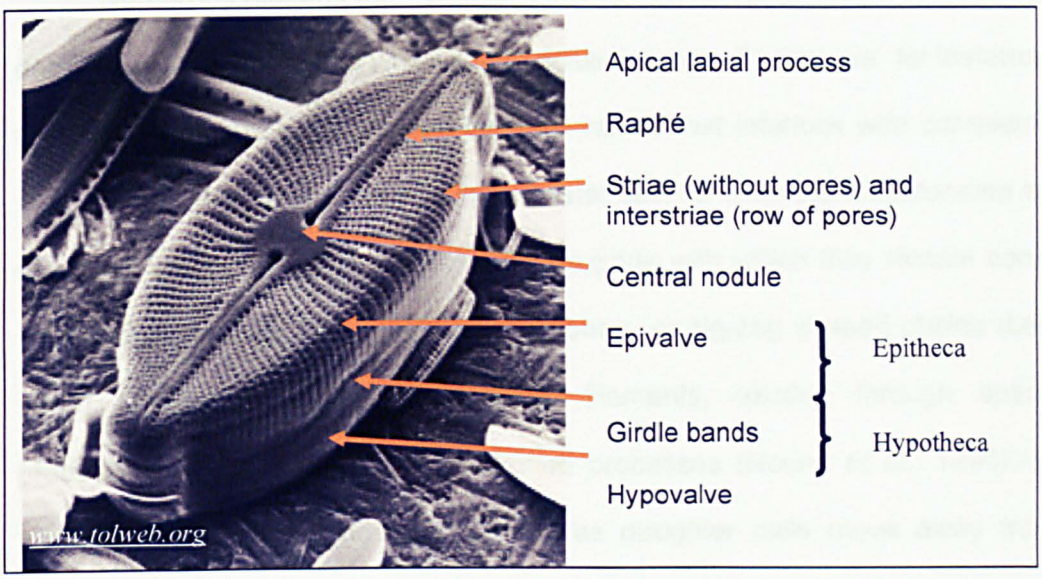


Figure 1.5: Scanning Electron Micrograph of a pennate diatom frustule.

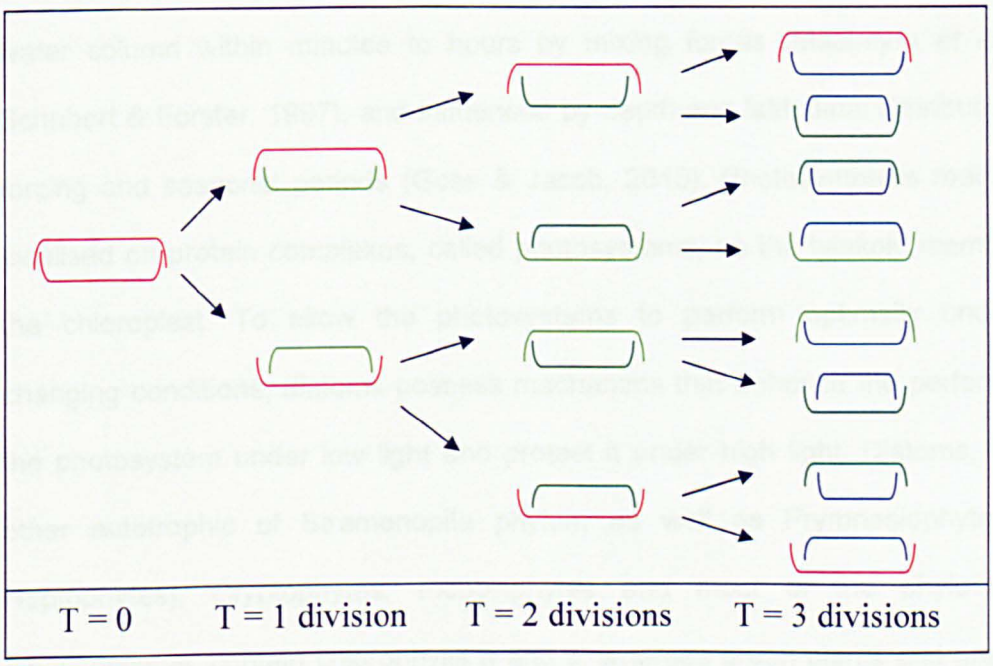


Figure 1.6: Cartoon of cell size reduction and division during vegetative growth.

Numerous diatoms form chains. Various mechanisms of chain formation exist, depending on the group to which the species belongs. Chains are, for instance, held together by means of extensions on their valves that interlock with complementary structures on the valve of their adjacent sister cell, for example *Skeletonema* species (Sarno *et al.*, 2005). Others form mucilage pads with which they remain connected with their sister cells thus giving rise to ribbon- or zig-zag shaped chains (Leppard, 1995). Thalassiosirales possess chitin filaments, exuded through specialized processes, called fultoportulae or strutted processes (Round *et al.*, 1990). Raphid pennates generally do not form chains as daughter cells move away from one another. However, there are exceptions and the stepped chains of *Pseudo-nitzschia* and the ribbons of *Fragilariopsis* are among them.

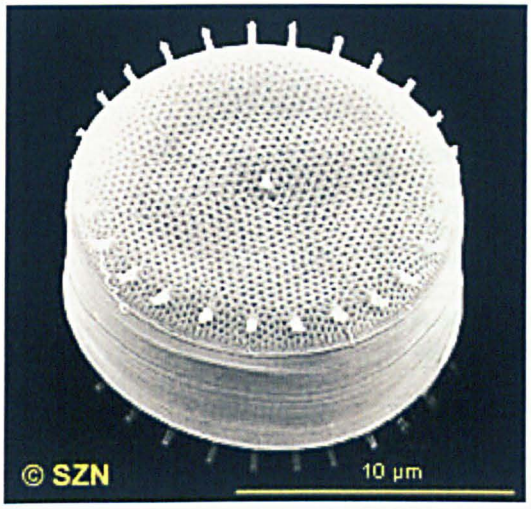
Phytoplankton have to adapt rapidly to irradiance changes driven through the water column within minutes to hours by mixing forces (MacIntyre *et al.*, 2000; Schubert & Forster, 1997), and influenced by depth and latitudinal distributions, tidal forcing and seasonal periods (Goss & Jacob, 2010). Photosynthesis reactions are localised on protein complexes, called photosystems, on the tylakoid membranes in the chloroplast. To allow the photosystems to perform optimally under these changing conditions, diatoms possess mechanims that enhance the performance of the photosystem under low light and protect it under high light. Diatoms, like most other autotrophic of Stramenopile phylum as well as Prymnesiophytes (a.k.a. Haptophytes), Cryptophytes, Picobiliphytes and most of the phylum of the dinoflagellates, contain chlorophylls a and c, whereas green plants and green algae possess chlorophylls a and b (Lavaud, 2007). Diatoms also share with most other autotrophic stramenopiles a range of photosynthetic antenna pigments that capture

light and transfer its energy to the chlorophyll; the most important of these is fucoxanthin (Lavaud, 2007). One of the photoprotective mechanisms adopted by the diatoms is related preferentially to the diadinoxanthin cycle (xanthophyll cycle: one de-epoxidation step). The surplus of energy, created under high light, is reemitted by chlorophyll fluorescence and is dissipated as heat by non-photochemical Chl fluorescence quenching (NPQ); the de-epoxidation of diadinoxanthin to diatoxanthin in the NPQ is induced by the acidification of the lumen caused by light-driven photosynthetic electron transport, and, prevents photoinhibitory damage (*e.g.* reactive oxygen species). Diatoms also possess the genes encoding the violaxanthin cycle (xanthophyll cycle: two de-epoxidation steps) (Coesel *et al.*, 2008), a cycle induced during rather long periods of high light exposure (Goss & Jacob, 2010).

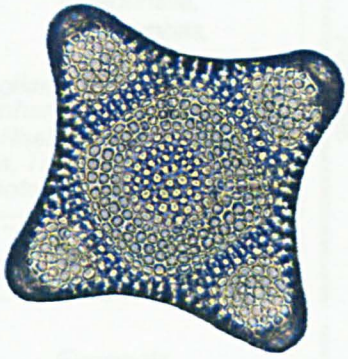
The possession of a large vacuole is thought to give diatoms a selective advantage over other Stramenopile microalgae. There is an absolute requirement for this vacuole during auxospore (*i.e.* key stage of the diatom cell formation after sexual reproduction) swelling where the rapid size increase is achieved primarily by means of vacuolar expansion (Mann, 1994). Yet, the vacuole also serves as a nutrient store in case the cell encounters a pulse of a particular nutrient (Kroth *et al.*, 2008). Diatoms can grow through up to four cell divisions without additional intake of a nutrient, such as phosphate, before the completely filled store runs out, depending of course, on the solubility of the particular nutrient, and hence, the storage capacity of the vacuole.

1.4.2. Diversity and taxonomy of the diatoms

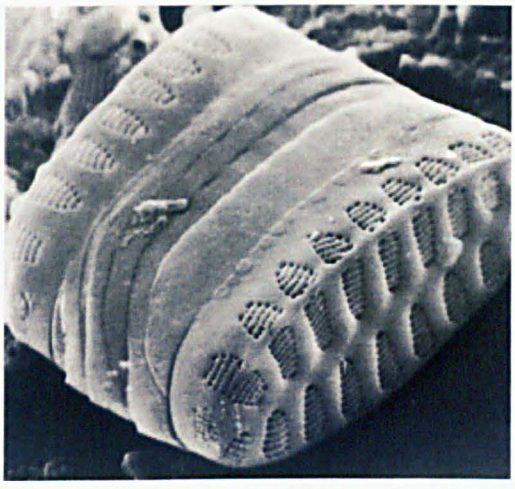
Diatoms can be divided taxonomically into radial centrics, multi- and bipolar centrics and pennates (Fig. 1.7). Radial centrics form the phylogenetically most ancient group whereas multipolar centrics constitute a lineage inside it (Fig. 1.8, (Kooistra, 2007)). In their turn, pennates form a lineage inside the multipolar centrics. Radial centric diatoms and Thalassiosirales possess valves that are circular in valve view, whereas multi- and bipolar centrics have valves that are multi-/bi-polar in valve view (Fig. 1.7). Pennate diatoms constitute a case of marked bipolarity with valves that are shaped like a boat (Fig. 1.7). The centric valve has a circular pattern with a centre from which striae (rows of pores) and interstriae (silica bars) radiate whereas the pennate valve possesses a bar-like pattern central midrib, from which striae and interstriae are oriented perpendicularly. Valves may possess additional structures depending on the taxonomic group to which the species belongs. Most centrics and many pennates possess rimoportulae, *i.e.*, processes with a lip-like (labiate) extension on the cell interior side. The majority of the pennates possess a mid-slit, called a raphe (Figs. 1.5 and 1.7), oriented alongside the midrib, which permits the cells to move actively. The raphe probably developed from apically located labiate processes in the a-raphid ancestry of the raphid pennate diatoms. In fact, raphid pennates form a clade or monophyletic group inside the groups of pennates without such a raphe (Fig. 1.8, (Kooistra, 2007)).



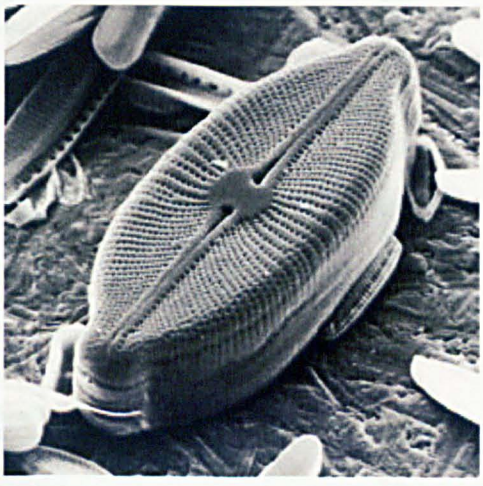
Radial centric diatom: www.szn.it
Thalassiosira allenii,
scanning electron microscopy



Multipolar centric diatom: www.diatomloir.eu
Amphitetras antediluviana,
light microsope



Araphid pennate diatom: www.tolweb.org
Punctastriata sp,
scanning electron microscopy



Raphid pennate diatom: www.tolweb.org
Cosmioneis sp,
scanning electron microscopy

Figure 1.7: Examples of frustules belonging to centric diatoms (A: radial centric order Coscinodiscophycidae, B: multipolar centric order Biddulphiophycideae) and to pennate diatoms (C: araphid pennate order Fragilariophycidae, D: raphid pennate order Bacillariophycidae).

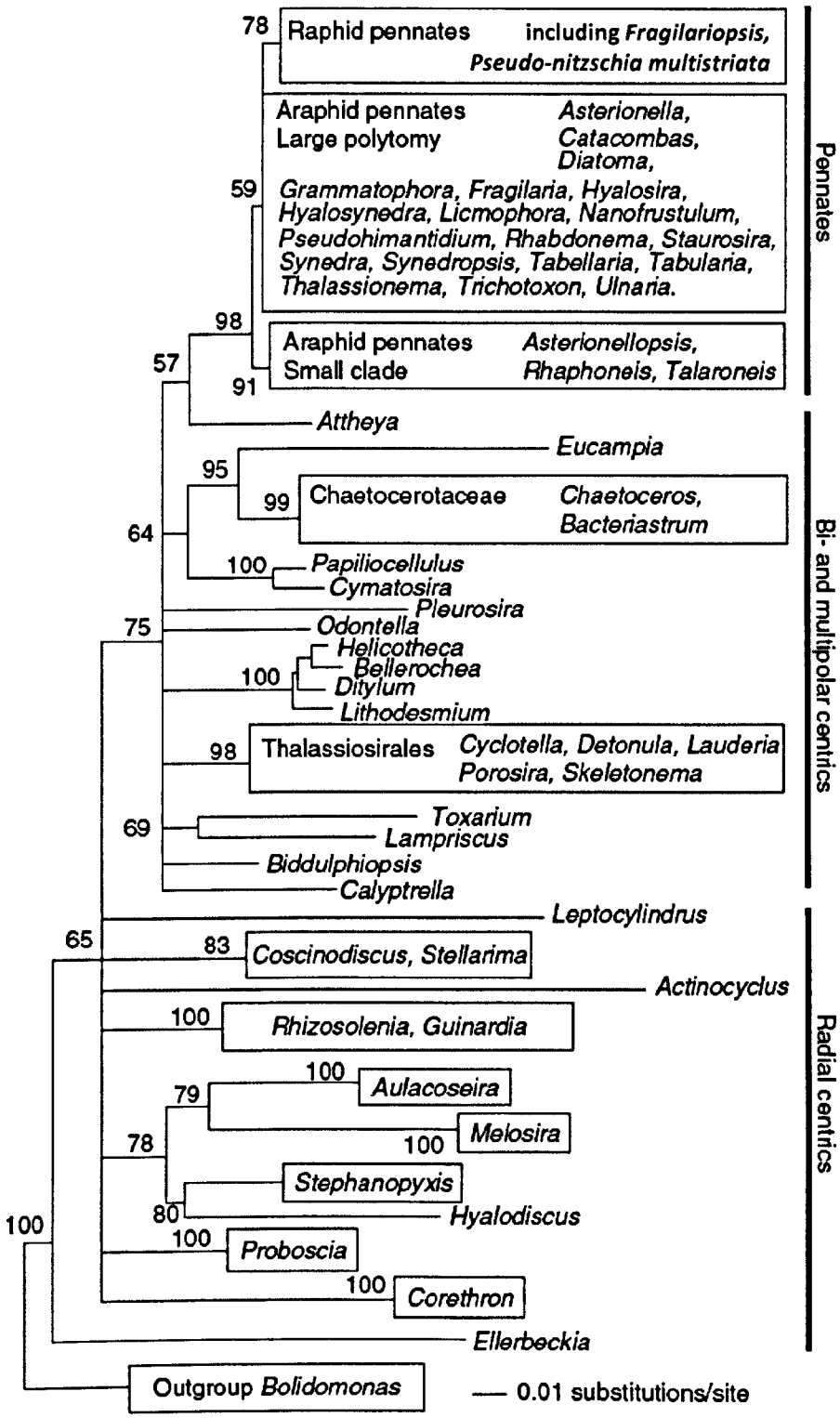


Figure 1.8 Diatom phylogeny adapted from Kooistra et al., (2007). Neighbor joining tree of maximum likelihood pair-wise distances amoung nuclear small sub-unit (SSU) ribosomal DNA sequences of diatoms.

The diatoms as a whole form a lineage inside the Stramenopiles. The latter, also called Heterokonta (Fig. 1.9), are defined by possessing a flagellar apparatus consisting of a backwards pointing, naked flagellum and a forwards pointing (*i.e.*, in the swimming direction) flagellum, ornamented with two rows of hairs, called mastigonemes. Obviously not all of the groups have flagella because of secondary losses. In diatoms, for instance, only the microgametes of the centrics possess a flagellum.

Within the Stramenopiles, the autotrophic lineages, including the diatoms, form a clade (monophyletic group) within a grade (a paraphyletic group) of non-photosythetic lineages. Other autotrophic lineages are, for instance, the brown macroalgae, the xanthophytes, the raphidophytes and the silicoflagellates (Fig. 1.9).

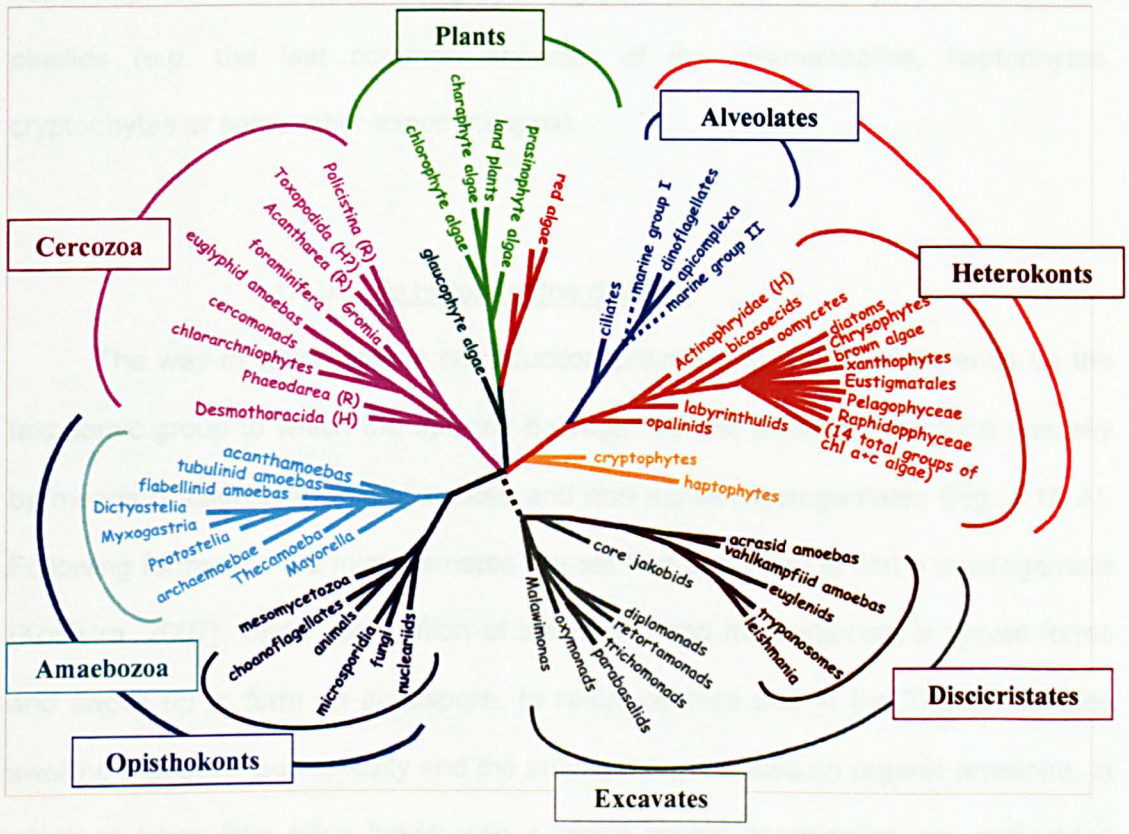


Figure 1.9: Eukaryote tree of life (Baldauf et al., 2004).

The plastids of the autotrophic lineages are characterized by possessing four membranes around their plastids through two successive endosymbiosis (Delwiche 1999). The abundance of membranes is due to successive capture and phagocytosis of an organism by another and the remodelling of the host genetic content (transfer of viable genetic material from the phagocytosed cell to the host). During the primary endosymbiosis a cyanobacterium was incorporated by a heterotroph. After cell modification the plastid is host-dependent and surrounded by two membranes. During the secondary endosymbiosis a red algal cell was incorporated into the ancestral host to give rise to a four-membrane plast envelope (the cyanobacteria ex-external membrane, the primary host food vacuolar and external membranes, and the secondary host food vacuolar membrane). The secondary host was either a heterotroph or from another lineage already possessing chlorophyll a and c in their plastids (*e.g.* the last common ancestor of the stramenopiles, haptophytes, cryptophytes or some other minor lineages).

1.4.3. Life history of the diatoms

The way in which sexual reproduction proceeds in a diatom depends on the taxonomic group to which the species belongs. Centric diatoms reproduce sexually by means of flagellated microgametes and non-motile macrogametes (Fig. 1.10 A). Following formation, the microgametes are set free and swim to find a macrogamete (Kooistra, 2007). Upon conjugation of the micro- and macrogamete a zygote forms and swells up to form an auxospore. In radial centrics and in the Thalassiosirales swelling proceeds isometrically and the auxospore generates an organic envelope, in which at times little silica flakes with a radial centric organization are embedded (Kooistra, 2007). Inside this envelope the new epi- and hypotheca form, thus giving

rise to the large initial cell. The only oddity of this initial cell is that the valves are round, forming a ball. When this initial cell divides along the equatorial plane, the first set of flattened valves is formed, reminiscent of Petri dishes. In the bipolar centrics, the swelling is shepherded into an anisometric (bi- or multipolar) shape by means of a series of properizonial bands, one formed after the other (Fig. 1.10 B). The new frustule is formed inside this auxospore.

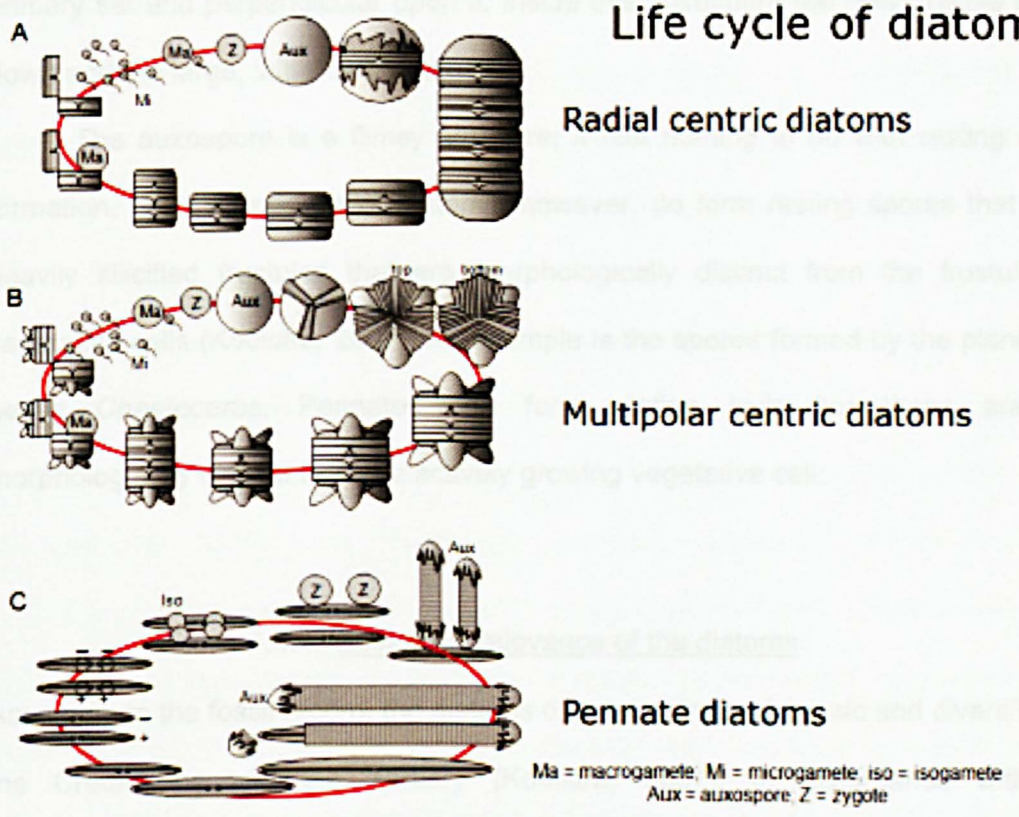


Figure 1.10: Cell cycle of centric diatoms (A,B) and pennate (C) diatoms.

In pennates, gamete formation proceeds differently. Two sexualized cells of opposite mating types align themselves alongside one another and perform meiosis together to form two gametes each (Fig. 1.10 C). Normally, meiosis would generate

four haploid gametes per diploid cell whereas pennate diatoms usually generate only two. It remains unclear where, how and why during the meiotic divisions half the nuclear material disappears (Round et al., 1990). The two gametes of one partner cell (gametangium) move in an amoeboid fashion toward the two gametes on the other partner and fuse with them (Kooistra, 2007). Upon gamete fusion, zygote swelling is constrained by a similar series of circular bands, called perizonial bands, into a cigar shape. Then a second series of elongated bands is formed inside the primary set and perpendicular upon it. Inside this auxospore the new frustule is laid down and the large, initial cell emerges.

The auxospore is a flimsy structure; it has nothing to do with resting spore formation. Certain lineages of diatoms, however, do form resting spores that have heavily silicified frustules that are morphologically distinct from the frustules of vegetative cells (Kooistra, 2007). An example is the spores formed by the planktonic genus *Chaetoceros*. Pennates can form resting cells but these are not morphologically distinct from the actively growing vegetative cell.

1.4.4. Ecological relevance of the diatoms

According to the fossil record, the diatoms originated in the Jurassic and diversified in the Cretaceous and the Tertiary (Kooistra, 2007). In accordance with the phylogenetic relationships (Fig. 1.8) among the extant representatives of the major taxonomic groups, the radial centrics appeared in abundance at *ca.* 105 Ma, followed by the multipolar centrics and the pennates at *ca.* 70 Ma. The youngest group, the raphid pennates, appeared first at around 55 Ma. The latter is now the most species-diverse group.

Diatoms are now ubiquitous in the photic zone of marine and freshwater ecosystems and in humid terrestrial environments (Kooistra *et al.*, 2007). Their cells can be found epiphytically, attached to rock and sand grains, or they drift above the substratum. The frustules of these benthic species are often heavily silicified. Many species occur in suspension in the plankton; these are generally more lightly silicified to reduce sinking rates (Kooistra *et al.*, 2009). Planktonic diatoms are common among the centrics whereas planktonic pennate lineages are rare. All planktonic pennates probably evolved from benthic epiphytic ancestry. Nevertheless, those few pennate lineages that are now found in the plankton, *i.e. Thalassionema* and relatives, *Asterionellopsis* and relatives, *Asterionella* and relatives, *Pseudo-nitzschia* and its probable relative *Fragilariopsis*, are each highly diverse.

The planktonic representatives of the diatoms in particular are important from a global ecological viewpoint. They are numerically most abundant because of the vastness of the planktonic habitat (Bracher *et al.*, 2009), cosmopolitan (Hasle, 2002; Kooistra *et al.*, 2008), even though most planktonic diatoms are confined to coastal and upwelling ecosystems and trapped in ocean currents/gyres. Together with other phytoplankton representatives, planktonic diatoms are also major net-fixers of CO_2 (Kooistra et al., 2007). A small part of plankton primary production in coastal regions will be buried in the accumulating sediment layers, to be transformed ultimately in petroleum, unlike virtually all the terrestrial primary production, which is decomposed back into CO_2. Apart from being important producers of organic matter, they also play a major role in the cycling of silica as they fix it to construct their frustule.

1.5. The genus *Pseudo-nitzschia*

1.5.1. Morphology and diversity of *Pseudo-nitzschia* species

Pseudo-nitzschia is a diverse genus of planktonic raphid pennate diatoms. Its defining character is that the cells can form chains to give stepped colonies. Following cell division, each daughter cell slides over the valve of its sister until the two cells touch only at their cell apices: in this way chains develop (see (D'Alelio *et al.*, 2009b)).

The genus *Pseudo-nitzschia*, first described by Peragallo & Peragallo (1900) was initially assigned as a section of *Nitzschia* because of their similar morphological, but phylogenetically distinct characteristics (Hasle, 1994), such as: i) potential sigmoidal frustule in girdle view (*e.g. N. sigmoidea* and *P. multistriata*), ii) linear in valve view, iii) silificated valves (heavier in *Nitzschia* than in *Pseudo-nitzschia* species), iv) presence of a raphe (unbroken for *Nitzschia* but eccentric and partly reduced for *Pseudo-nitzschia* species), v) striated valves (almost uniseriate striae for *Nitzschia* species), the valves ends attenuated (gradually attenuated in *Pseudo-nitzschia* species) and vi) the ability to glide (Hasle, 1994).

Pseudo-nitzschia is probably sister to *Fragilariopsis*, a planktonic genus of generally robust frustules, though with large pores in the striae to render the frustules both lightweight and architecturally strong. *Fragilariopsis* is predominantly known from cold water. Both genera are related to the benthic genus *Nitzschia*.

1.5.2. Societal relevance of *Pseudo-nitzschia species*

The prime reason why *Pseudo-nitzschia* has been the subject of considerable research interest is that many of its species have been shown to be able to produce the toxin domoic acid (DA, Fig. 1.11). This heterocyclic amino acid is a potent neurotoxin and the causative agent of amnesic shellfish poisoning (ASP) in humans and other vertebrates that have ingested shellfish or other filter feeders in which the toxin has accumulated. As of now, the following species are known to produce DA: *P. multiseries* (Hasle) Hasle, *P. delicatissima* (Cleve) Heiden, *P. pseudodelicatissima* (Hasle) Hasle, *P. australis* Frenguelli, *P. seriata* (Cleve) H. Peragallo, *P. fraudulenta* (Cleve) Hasle, *P. turgida* (Hustedt) Hasle, *P. pungens* (Grunow ex Cleve) Hasle and *P. multistriata* (Takano) Takano; (Bates, 2000). Outside *Pseudo-nitzschia* two other species are known to produce DA, namely *Amphora coffeaformis* (Agardh) Kützing and *Nitzschia navis-varingica* Lunholm & Moestrup; (Bates, 2000). Many other *Pseudo-nitzschia* species exist, but none of these have been demonstrated to be toxic. The species that forms the subject of the current study is *P. multistriata,* which is mildly toxic. Toxicity levels are generally low during rapid growth, but increase markedly when cells run out of phosphate or silica (and possibly other essential nutrients). N-shortage generally does not boost DA production, probably because nitrogen is needed for the production of amino acids such as DA. Monitoring for toxic blooms of *Pseudo-nitzschia* therefore includes monitoring of the species composition of populations by light microscopy and monitoring for the toxin.

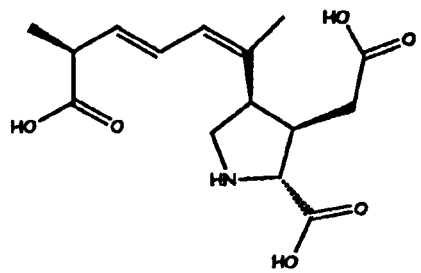


Figure 1.11: Domoic acid molecule (www.nwfsc.noaa.gov)

Scholin and colleagues (Miller & Scholin, 1998; Miller & Scholin, 2000; Scholin *et al.*, 1999) developed species-specific molecular probes targeting the Large Subunit (LSU) of the ribosomal DNA gene for use in Fluorescence In-Situ Hybridisation (FISH) methods to detect toxic *Pseudo-nitzschia* species from plankton samples. The method worked well in the geographical regions where it had been developed and identified the local toxic species. However, researchers deploying the probes in other regions discovered that they did not work properly: the reason was that the species diversity was much higher than initially assumed with genetically distinct, but morphologically similar geographically separated populations. An alternative to FISH methods was the FITC-lectin binding assays (Cho *et al.*, 2002; Rhodes *et al.*, 2000; Rhodes *et al.*, 1998), that allowed the discrimination between morphospecies from the same locality. However, again, different patterns were observed for the same morphospecies from different localities.

The level of toxicity differs not only among species but also among strains within a species. Results of Amato *et al.* (2010) revealed that about half the strains sampled from a *P. multistriata* bloom showed no detectable toxicity at the end of the exponential growth in culture, *i.e.*, at the moment when toxicity in the cells usually

peaks. Of those that showed toxicity, the concentrations differed from 0.1 to 282.5 fg.cell^{-1}. Even the offspring of toxic parents showed significant differences in toxin concentration with about half of the F1 strains not showing any toxicity. The ability to produce DA was strain specific because when tested throughout their vegetative propagation, a toxic strain remained capable of producing DA throughout whereas a non-toxic strain remained non-toxic.

The reason why *Pseudo-nitzschia* species produce DA is not understood. Many microalgae produce secondary metabolites that function to make life of their competitors or natural enemies miserable. Such metabolites are called allelopathic compounds. Examples from terrestrial habitats are phenols (inhibition of mineral absorption), quinones (gene expression regulation) and terpens (vegetal growth inhibition; (Tillmann *et al.*, 2008)). Examples of allelopathic substances in aquatic environments are amino acid analogues, purine derivatives, cyclic imines or non-nitrogen compounds (Cembella, 2003). Such substances can modify the predator behaviour and morphology or induce a feeding inhibition (allomones). Even more elaborate defence mechanisms abound as microalgae can exude compounds called synomones to attract the predator's predator (Cembella, 2003). The problem with DA is that it has never been shown to harm grazers such as copepods or filter feeders such as mussels and oysters (Cembella, 2003). Copepods accumulate DA in the lysosomes of their gut cells and in extracellular granules beneath the cuticule and the basal lamina before they excrete the toxin in faecal pellets (Barka, 2007). Shellfish and small benthic fish accumulate DA in their tissues and viscera, but without any ill effect (Busse *et al.*, 2005; Vigilant & Silver, 2007). Cephalopods eat shellfish, but these organisms do not only accumulate DA, they also transform it into less toxic derivatives and then excrete it through their digestive glands (in direct contact with

the food eaten), their kidney and more particularly, their branchial hearts (Costa *et al.*, 2005).

Domoic acid might be produced for entirely different reasons. Notably, levels are high under nutrient stress. Moreover, DA is kept inside the cell. This is not what would be expected from an allelopathic compound. Instead, DA has been identified as a potential metal chelator in the marine environment. Such a chelator may modify the availability of metal ions such as iron and copper (Rue & Bruland, 2001), or play a role in copper detoxification.

Whatever the function of these secondary compounds, they can create collateral damage when they are accumulated higher up in the food web. Mammals and birds do not have detoxification and excretion mechanism described above, and DA acts as an analogue of the neurotransmitter glutamate inducing the apoptosis of neuronal cells (Carvalho Pinto-Silva *et al.*, 2008), including the brain (Meldrum, 2000).

Apart from DA, which causes ASP, several other examples of such effects exist. Examples are: paralytic shellfish poisoning (PSP, due to saxitoxin), ciguatera shellfish poisoning (due to polyether), diarrhoeic shellfish poisoning (DSP, due to okadaic acid) or neurotoxic shellfish poisoning (NSP, due to brevetoxins) (Cembella, 2003; Zingone & Enevoldsen, 2000). The production of toxins may modify community structure, transduction energy and the pattern of the food web by inducing negative impact on mammals (illness, death), on marine ecosystems (anoxia) and on resources (fisheries, mariculture, recreation) (Zingone & Enevoldsen, 2000). The algal blooms that cause these phenomena are collectively known as harmful algal blooms (HABs).

1.6. The species *Pseudo-nitzschia multistriata*

1.6.1. Morphology of *Pseudo-nitzschia multistriata*

The species *P. multistriata* can be distinguished from other *Pseudo-nitzschia* species under light microscope (Fig. 1.12) because of its pronounce sigmoid shape in girdle view. In valve view its shape is linear with pointed ends, the raphe is present but the central interspace is missing. Cells are between 38-50 μm long and 2.8-3.8 μm wide. The valve possesses 37 to 46 interstriae, composed of two rows of circular poroids, and 23 to 28 fibulae in 10 μm (Sarno & Dahlman, 2000).

According to distribution ranges illustrated in Hasle (2002), *Pseudo-nitzschia multistriata* is a temperate species that occurs closest towards the tropics, and it is observed generally in the warmer temperate waters. *Pseudo-nitzschia multistriata* was first identified in Japan (Orlova *et al.*, 2008; Takano, 1993; 1995), in New Zealand from October to January (southern spring and summer) (Rhodes et al., 2000) and Korea in autumn (Cho *et al.* 2002). It appeared recently at monitoring stations in Europe, both along the Atlantic coasts (Portugal, Spain, France) (Churro *et al.*, 2009; Nezan & Chomerat, 2007; Orive *et al.*, 2010), Mediterranean coasts (Spain, Italy, Greece) (Moschandreou & Nikolaidis, 2010; Quijano-Scheggia *et al.*, 2008a; 2008b; 2006; Sarno & Dahlman, 2000; Zingone *et al.*, 2006) and the Moroccan coasts (Akallal *et al.*, 2002) (Fig 1.13). The species has also been observed in South African waters and in the Gulf of Mexico (personal communications). The recent apparent range expansion of *P. multistriata* may result from ballast water discharges, from natural migration linked to climate change or alternatively, it may have always been present, but below the threshold for detection.

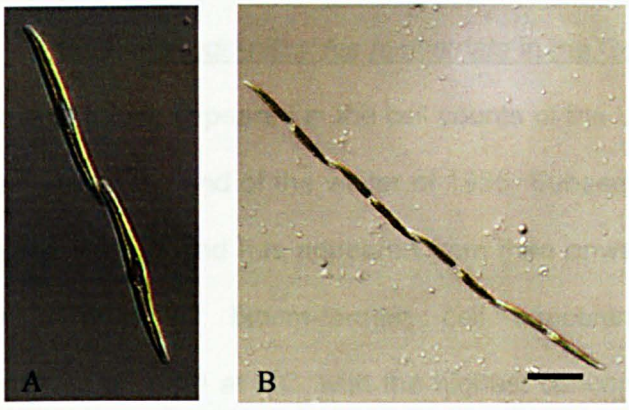


Figure 1.12: Pseudo-nitzschia multistriata *in culture (A,B: girdle band view, light microscope AXIOPHOT x10x40x1.25, image ZEISS AXIOCAM, resolution 2600x2060, AXIOVISION 3.1 software, scale bar 20 µm).*

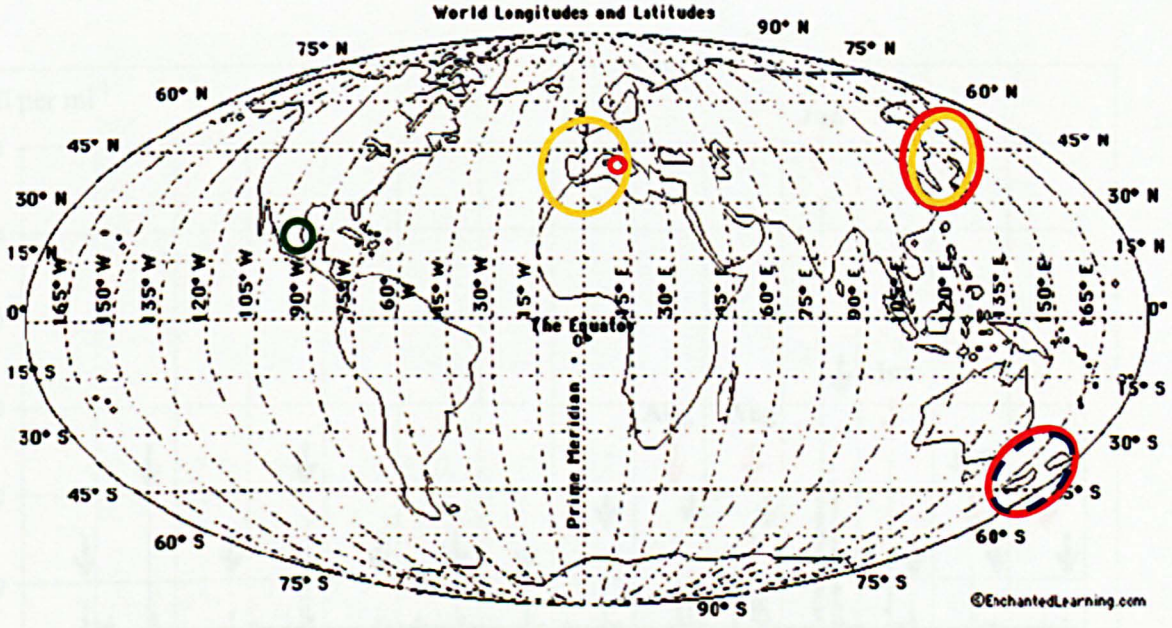


Figure 1.13: Location of Pseudo-nitzschia multistriata *blooms in occurrence and their seasonality. Main occurrences are illustrated with a solid line and minor occurrences with a dashed line. Occurrence in summer (orange), autumn (yellow), winter (blue) and unknown (green)). Data have been gathered from the following publications: (Cho et al., 2002; D'Alelio et al., 2010; Nezan & Chomerat, 2007; Orive et al., 2010; Orlova et al., 2008; Quijano-Scheggia et al., 2008a; 2008b; 2006; Rhodes et al., 2000; Takano, 1993).*

1.6.2. *Pseudo-nitzschia multistriata* in the GoN

Pseudo-nitzschia multistriata appeared in the cell counts of the LTER-MC for the first time in a sample taken at the end of the winter of 1995. Subsequently it appeared in the autumn-samples of 1996 and has appeared from then onwards every year. The species is not a dominant bloom-former; cell concentrations remain low (max: 6.9 x10^5 cell·l^{-1}; Fig. 1.14) at MC, with the highest concentrations occurring in autumn and tailing off into the winter of the next year (Figs. 1.14 and 1.15). The summer bloom shifted from August in the years 2004 and 2005, to June-July in the years from 2006 to 2009 (Fig. 1.14). Figure 1.15 shows cell concentration of *P. multistriata* over the period 1995-2006.

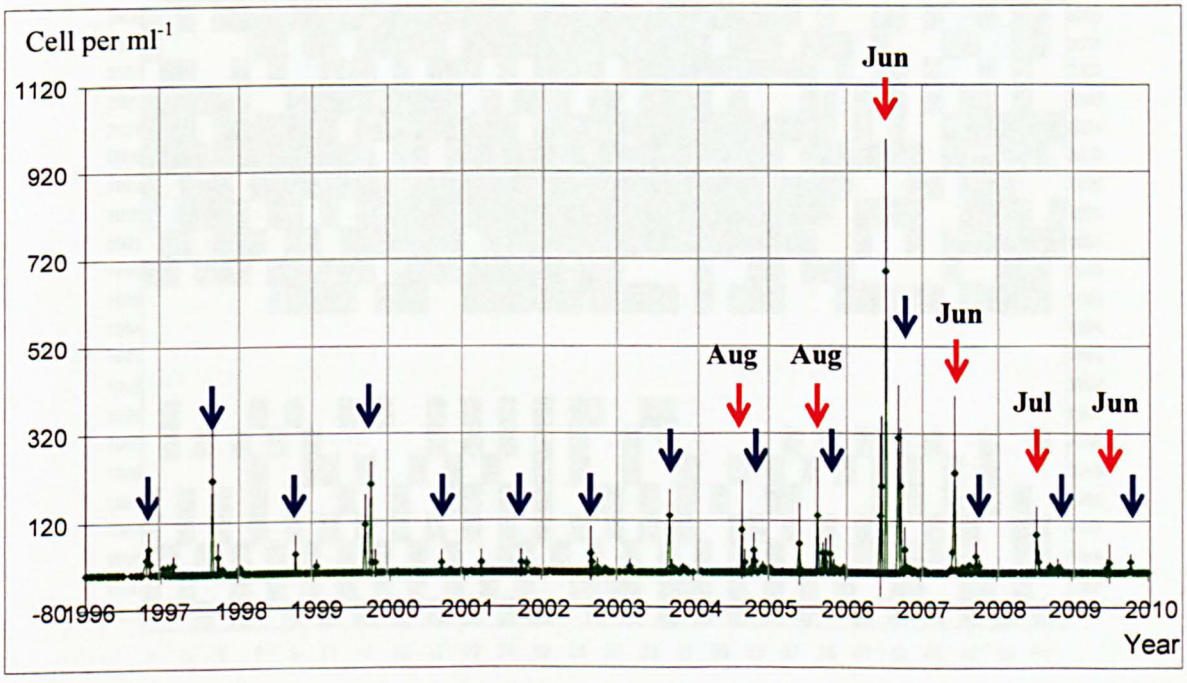


Figure 1.14: Occurrence of P. multistriata *at LTER-MC station (1995-2009) from weekly sampling (data provided by Dr D. Sarno): blue arrow: autumn bloom, red arrow: early-summer bloom (Aug: August, Jun: June, Jul: July).*

1.6.3. A biennial life cycle of *Pseudo-nitzschia multistriata*

The life cycle of *P. multistriata* appears to span two years. A cohort, i.e., all cells issued from a same sexual event and all their clones, starts at a cell length of *ca.* 80 μm, but does not show up in the environmental monitoring of that year because these cells are rare. In the next year, the cohort shows up as the largest one among two distinct cell size classes; the smallest of the two representing the members of the parental cohort that have, for whatever reasons, not reproduced sexually, that are now no longer sexually inducible and that are destined to die. The subsequent year the cohort is well within the size range of sexual maturity (between 55 μm and 30 μm) and towards the end of the bloom its members can reproduce sexually.

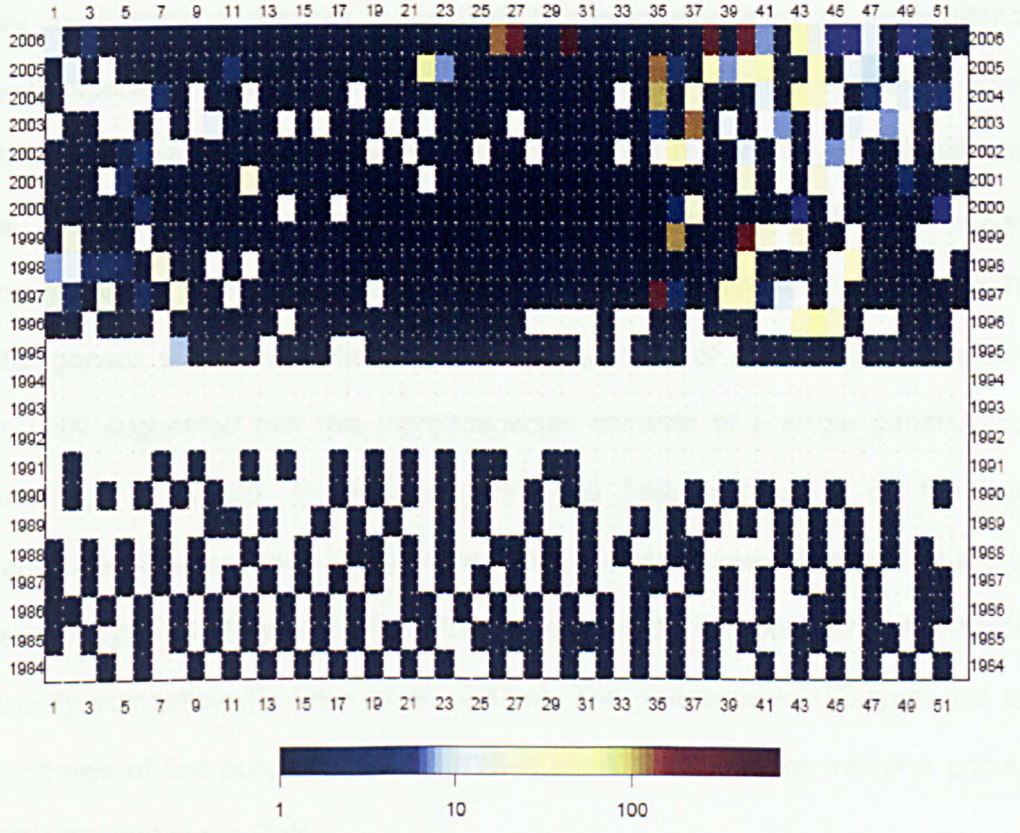


Figure 1.15: Cell concentration of Pseudo-nitzschia multistriata *at LTER-MC station from 1984 to 2008 (data provided by Dr D. Sarno and by D'Alelio et al., 2010). Colour*

legend depicts cell concentration (cells·ml^{-1}) in a logarithmic fashion. Numbers along horizontal axis denote week numbers (1-52) over the years 1984-2006 (see vertical axis). White rectangles denote 'no sample taken'.

This cohort is the only one visible, showing up as a single size class in the counts. The third year, this cohort has now become the smaller one of the two distinct size classes of that year, is not longer sexually inducible and is destined to die out. The result of such a biennial life cycle is that the population blooms every year but shows sexual reproduction only every odd year (D'Alelio et al., 2010).

1.6.4. Differences in the ITS sequences among strains of *Pseudo-nitzschia multistriata* (D'Alelio *et al.*, 2009a)

Pseudo-nitzschia multistriata in the GoN has been examined for the existence of cryptic species. In fact, many initially morphologically defined species in *Pseudo-nitzschia* consist of multiple, genetically distinct and reproductively isolated species that are difficult or impossible to distinguish morphologically. Examples are the morphospecies *P. delicatissima* and *P. pseudodelicatissima* (Amato *et al.*, 2007). An initial genetic survey of partial LSU rDNA sequences of *P. multistriata* strains from the GoN suggested that this morphospecies consists of a single genetic species (Orsini *et al.*, 2002). Subsequent work has been conducted on the Internal Transcribed Spacer region (ITS) of the rDNA. Results showed two distinct ITS-types (called A-type and B-type), with a mixed type (called A/B-type) and all ITS-types are sexually compatible (D'Alelio et al., 2009a). The difference in ITS-types led to the hypothesis of two pure populations (ITS-A and ITS-B) merging into one population (ITS-A/B-type) in the GoN.

1.7. The research questions and goals of the thesis.

Pseudo-nitzschia multistriata is probably a coastal species growing within the semi-enclosed GoN, therefore, one might expect that the genetic structure of the population of *P. multistriata* collected at a station in the GoN stays homogenous over the course of a year. However, currents in the GoN, the flanking Gulfs and Tyrrhenian Sea, reveal a clear distinction between coastal eutrophic waters and offshore oceanic oligotrophic water, and as well as a clear cross-influence of coastal waters from the three gulfs. Part of a coastal water mass in which the species grows can be transported out into the open sea through eddies. Even if the bulk of the coastal water in the GoN remains protected by the semi-enclosed nature of the Gulf, there may be migration between different regions in the Mediterranean Sea, especially over short distances, as demonstrated by the water masses flowing from the GoG in the GoN and from the GoN into the GoS. Therefore, several regional populations of *P. multistriata* may exist along the coastlines of the Mediterranean, forming parts of a semi-fragmented metapopulation (*i.e.*, one species whose populations are spatially separated).

Winter storms may flush out the system and replace a resident population with one that is assembled from a mix of localities. This would result in an absence of genotypic differences between the summer and autumn populations of one year, but marked differences between the local populations in different years.

The alternative hypotheses are then that there is just one thoroughly mixed Mediterranean population, with minor gene flow from elsewhere, or a single, resilient, local population (Italian west coast including GoG-GoN-GoS) with minor gene flow from elsewhere. Although the geographic scale differs between the two hypotheses,

the resulting population structure would be highly similar and persistent over the years.

Goals of the present study were to determine the genetic structure of *P. multistriata* population(s) in the GoN at the LTER-MC station using polymorphic genetic loci (*i.e.* ITS and microsatellites) over blooms, cohorts, time and life cycle.

- How to maintain a significant number of *P. multistriata* strains in culture from the sampling to population analyses? (Chapter 3).

- How do ITS patterns evolve over seasons (early-summer and autumn) and successive years (2008-2009)? Is the variability of ITS patterns conform to previous studies when the sampling effort is increased? Is the theory of two populations merging in the GoN confirmed over the 2008 and 2009 blooms? How variable is the ITS-A/B pattern? (Chapter 4).

- Are the seven microsatellites developed for *P. multistriata* polymorphic, and can they assess all population genetic parameters? (Chapter 5).

- What is the population genetic structure of *P. multistriata* in the GoN over the 2008 and 2009 blooms? Does the population genetic structure change over seasons (early-summer and autumn) and over cohorts? (Chapter 6).

- The ITS polymorphism in *P. multistriata* suggests the existence of multiple populations in the GoN. Do the microsatellites show the same phenomenon? What is the bloom dynamic observed in the field as revealed by comparisions of ITS and microsatellite patterns, and cohort succession? (Chapter 6).

- Are the two populations of *P. multistriata* as defined with the microsatellite markers sexually isolated? Does *P. multistriata* genetic structure change after a sexual event? (Chapter 7).

- Does the genotype of *P. multistriata* strains change during the vegetative phase of the life cycle? (Chapter 8).

1.8. Background on procedures and methodologies used: why these particular ones?

ITS markers permit the identification of interindividual variation and are used in species phylogeny analyses (Montresor *et al.*, 2003). However, Evans *et al.* (2007) state that for some species ITS is not a suitable marker to establish a diatom barcode. A good marker to study population genetic structure needs to be neutral (being placed in a non-coding region), highly variable, co-dominant (to estimate both genotype and allele frequencies), bi-parentally inherited and subject to recombination during sexual reproduction (Ouborg *et al.*, 1999). The techniques potentially at our disposal are: restriction fragment length polymorphism (RFLP), single strand conformation polymorphism (SSCP), single nucleotide polymorphism (SNP), variable number of tandem repeats (VNTR: microsatellites) and amplified fragment length polymorphism (AFLP).

RFLP and SSCP techniques are suitable for DNA sequence variation identifications (Féral, 2002; Ouborg *et al.*, 1999). RFLP is a technique that qualifies the gain or loss of mutation after enzymatic digestion of the whole genomic DNA. SSCP is a technique that allows the detection of small changes in single stranded DNA (mutations, insertions, polymorphisms). Genomic DNA is firstly digested by

endonucleases and then denatured in an alkaline solution. The products migrate through a polyacrylamide gel, are then transferred onto a nylon membrane and hybridized with a specific probe. The variability of DNA sequences (double or single stranded) influences the conformation of the strand and its migration through the gel. It is an easy, cheap and potent technique, but acrylamide is toxic in the non-polymerized form and its use is therefore no longer recommended.

AFLP, SNP and VNTR analyses are used to identify DNA polymorphism (Féral, 2002). AFLP is a qualitative analysis of the presence/absence of restriction fragments. The genomic DNA is digested by restriction enzymes and complementary adaptors are ligated to each end of the double-stranded and cut DNA. Then adaptor-specific primers amplify the cut DNA fragments, and fragment length patterns are compared after migration through an electrophoresis gel. This method has the disadvantage of not being species-specific, therefore it is impacted by the presence of DNA isolated from any contaminating organisms in culture: it is also less variable than VNTR analysis (Féral, 2002). The SNP technique allows the differentiation of genomic DNA on the basis of single base pair variation, which can be used both on coding and non coding regions of genes, and intergenic regions. It is a powerful method but the variability of one base pair and the error of machine reading and misinterpretation are higher than other methods. VNTR is a technique that relies on the identification of a core of tandemly repeated nucleotide sequences, such as minisatellite-markers, which possess a core of 10 to 100 bp, and microsatellite markers which are smaller, with a repeating core of few to 10 bp. Microsatellites have the advantages of being neutral and codominant markers, to have a high mutation rate (*ca.* 10^{-3} to 10^{-5}) (Hancock, 1999; Ouborg *et al.*, 1999), to be linked to hypervariable loci, being analysable using automated techniques (Procaccini &

Maltagliati, 2004) and being taxon specific. They have been used in population genetic structure to measure gene flow among populations and to identify the structure of metapopulations (Adams *et al.*, 2009; Evans *et al.*, 2004; Evans *et al.*, 2005; Ouborg *et al.*, 1999; Rynearson & Armbrust, 2000; Wattier *et al.*, 1997).

CHAPTER TWO

MATERIALS AND METHODS

2.1. Introduction

In the Chapter Materials and Methods, the different techniques used in the Ph.D.-thesis (Chapters 3 to 8) are explained: *Pseudo-nitzschia multistriata* strains isolation and culture, DNA extraction methods (manually and robotically -optimized), and details on DNA-purity and -quality techniques (nanodrop and agarose gels). Then, microsatellite- and ITS protocols used (amplification, purification, sequencer analysis) are presented.

2.2. Culture of *Pseudo-nitzschia multistriata*

Two types of field samples were collected weekly from the Long-Term Ecological Research station MareChiara (LTER-MC): at the water surface (horizontal net sample) and from 60 meters to the water surface (vertical net sample). The plankton net used for the horizontal track was an APTEIN, 100 cm length, diameter 40 cm and mesh size 20 μm, (Ageotec S.r.l, Zola Predosa, Italy). Directly upon arrival at the Stazione Zoologica Anton Dohrn (SZN), some droplets from the bottom of the sampling bottle were diluted in sterile seawater in a Petri dish to allow the recognition and isolation of the individual cells or chains of *Pseudo-nitzschia multistriata*. A chain of cells consists of genetically identical individuals derived as the vegetative progeny of a single cell. Cells, either solitary or in chains were identified under an inverted light microscope (100x magnification; Leica Microsystems DMIL, Wetzlar, Germany).

Table 2.1 F/2 Medium composition (Guillard, 1975; Guillard & Ryther, 1962). List of components (Bases, Trace metals and Vitamin solutions) and their concentration in the final F/2 medium. V_I indicates the initial volume in which components are added. V_F indicates the final volume of F/2 medium.

	Components	Concentration in Final Medium (M)
Bases	$NaNO_3$	8.82×10^{-4}
	NaH_2PO_4 , H_2O	3.62×10^{-5}
	Na_2SiO_3 , $9H_2O$	1.06×10^{-4}
Trace metals solution	$FeCl_3$, $6H_2O$	1.17×10^{-5}
	Na_2EDTA, $2H_2O$	1.17×10^{-5}
	$MnCl_2$, $4H_2O$	9.10×10^{-7}
	$ZnSO_4$, $7H_2O$	7.65×10^{-8}
	$CoCl_2$, $6H_2O$	4.20×10^{-8}
	$CuSO_4$, $5H_2O$	3.93×10^{-8}
	Na_2MoO_4, $2H_2O$	2.60×10^{-8}
Vitamin solution	Thiamine, HCl (vitamin B_1)	2.96×10^{-7}
	Biotin (vitamin H)	2.05×10^{-9}
	Cyanocobalamin (vitamin B_{12})	3.69×10^{-10}
V_I	Filtered natural seawater	950 ml
V_F	Filtered natural seawater	Up to 1 litre

Individual cells or chains were isolated by capillarity with a micro-glass pipette and cleaned by washing them through several droplets of F/2 (Table 2.1, Sigma-Aldrich S.r.l., Milan, Italy) medium until there were no other organisms visible in the droplet anymore. Each cell or chain was placed in a separate well of a 12-well-culture-plate (Costar® 3513, Corning Incorporated, NY, USA) containing 2 ml F/2

filtered medium and incubated at 22 °C, *ca.* 80 μmol photon·m^{-2}·s^{-1}, photoperiod 12:12. After approximately four days plates were examined to determine the success of the isolation, and the purity of the cell cultures developing from the isolated cells. Each successful unialgal culture was assigned a unique strain number. After approximately one week cultures reached a cell concentration about 10^3 cell·ml^{-1} (corresponding to about 30 ng purified DNA·μl^{-1}) and was ready to be transferred to a larger volume of medium, or to be processed for robotic DNA extraction. *P. multistriata* strains are defined as the clonal vegetative progeny of a single cell or chain of cells, and cultures are defined as a volume of growth medium containing cells.

Genomic DNA was extracted from a culture either manually or robotically. For manual extraction the 2 ml starter cultures were transferred into 25 ml F/2 medium in a 25 ml-bottle (Corning® Flask, code 430639, Corning Incorporated) for further growth (*ca.* 4 days). For robotic extraction the 2 ml of culture was transferred directly to the extraction plate and 2 ml of fresh medium was added to the culture well.

2.3. Extraction of DNA from cultures (protocols, tests and limits)

All data presented in Chapters 3 to 8 have been derived from manually extracted DNA. Manual extraction (40 extractions per run instead of 96) was the technique of preference because despite being far more time consuming, it was the cheapest, most reliable in providing purified DNA and with less technical (robotic error) and genetic error inducement (Whole genome amplification (WGA)). The following paragraphs provide details on the manual extraction (2.3.1.) and robotic extraction (2.3.2.) procedures.

2.3.1. Manual DNA extraction.

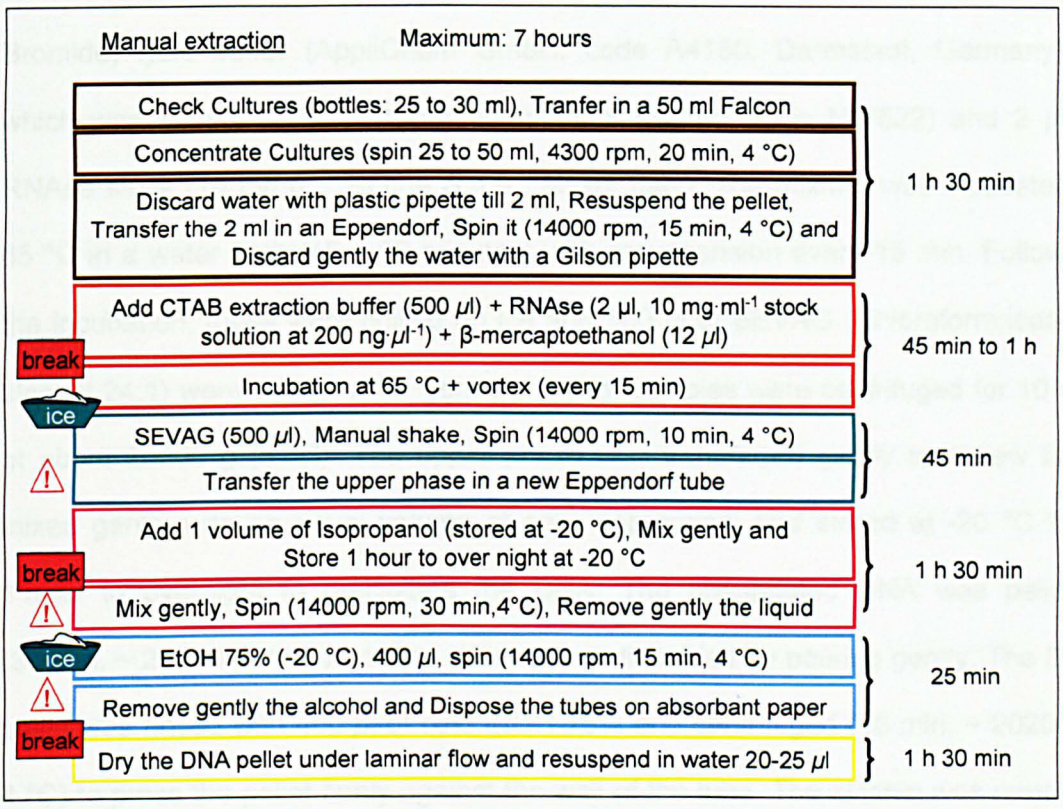


Figure 2.1: Outline of the manual protocol developed to obtain a high quality of DNA from manual extraction (CTAB extraction buffer) and culture centrifugation from Pseudo-nitzschia multistriata *species.*

The manual extraction protocol is summarised in Fig. 2.1. Once reaching about 10^3 cell·ml^{-1}, 25 ml of culture were harvested by centrifugation for 25 min at 3928 g (*i.e.* 4300 rpm, see Fig. 2.1) at 4 °C in 50 ml centrifuge tube (code 430828, Corning Incorporated). The supernatant was removed gently with a sterile plastic pipette to leave a fluffy pellet in about a 2 ml volume. The pellet was re-suspended, transferred into a 2 ml Eppendorf tube (Eppendorf S.r.l., Milan, Italy) and centrifuged for 15 min at about 20200 g (*i.e.* 14000 rpm, see Fig. 2.1) and at 4 °C. After the centrifugation

the supernatant was removed gently using a 100-1000 µl pipette Gilson (Gilson Inc, Middleton, WI, USA).

The pellet was resuspended in 500 µl of 2x CTAB (Cetyl Trimethyl Ammonium Bromide) lysis buffer (AppliChem GmbH, code A4150, Darmstadt, Germany) to which was added 12 µl of β-mercaptoethanol (Sigma, code M-7522) and 2 µl of RNAse stock (10 $mg \cdot \mu l^{-1}$, Roche S.p.A., Milan, Italy). The mixture was incubated at 65 °C in a water bath (45 – 60 minutes) with resuspension every 15 min. Following the incubation, tubes were chilled on ice and 500 µl of SEVAG (Chloroform:Isoamyl alcohol 24:1) were added. After gentle shaking, samples were centrifuged for 10 min at about 20200 g (4 °C). The upper phase was transferred gently to a new tube, mixed gently with an equal volume of cold isopropanol and stored at -20 °C from 1 hour to overnight to precipitate the DNA. The precipitated DNA was pelleted (30 min, ~ 20200 g, 4 °C) and the supernatant discarded by pouring gently. The DNA pellet was rinsed with 400 µl of cold EtOH 75% and centrifuged (15 min, ~ 20200 g, 4 °C) to press the pellet firmly against the wall of the tube. The alcohol was removed gently by pouring and the pellet was air-dried in a laminar flow hood. The DNA pellet was dissolved in 25 µl of H_2O overnight at 4 °C before use. Each DNA sample was diluted to 25 $ng \cdot \mu l^{-1}$ and 20 µl of the DNA solution was added to a well in a 96-well-plate (Unithermal full skirted plates 96 x 0.2 ml, Delchimica scientific glassware, code 20-5010310, Naples, Italy).

2.3.2. Robotic extraction.

2.3.2.1. Genomic DNA automated extraction

The automated extraction protocol was based on the Nucleospin 96 plant kit (Macherey-Nagel GmbH, Düren, Germany), and was carried out in collaboration with Drs Marco Borra and Elio Biffali of the SZN Molecular Biology Service. Samples (2 ml of a full 12-well-culture-plate with 40933 ± 16163 cells·ml^{-1}) were collected in a 96-deep-well-plate (Millipore, Billerica, MA, USA) and centrifuged (60 min, 3200 rpm, 4 °C). The supernatant was discarded and the pellets were frozen immediately at -20 °C. Pellets were re-suspended in an extraction buffer according to the kit manufacture's instructions.

DNA was extracted from resuspended cells samples with CTAB lysis buffers specially developed for plants, which contains denaturing agents, detergents and chaotropic salts. The latter disrupt non-covalent bonds such as hydrogen bonds in proteins, thereby denaturing proteins, and at high concentration also RNA and DNA. The lysis mixture was cleared by centrifugation (20 min, 6000 g) to remove polysaccharides, contaminants and residual cellular debris. The clear filtrate was then mixed with binding buffer and ethanol to create conditions for optimal binding of the DNA to the silica membrane, and loaded onto the binding plate containing 96 independent columns. The silica membrane was washed with two different buffers and dried by vaccum (300 sec, 846.59 hPa). The DNA was eluted by vaccum in low salt buffer and was ready-to-use for downstream applications. A gripper tool was provided for automated disassembly and reassembly of the stack. The genomic DNA

extraction method on Biomek® FX purified a single 96-well plate of samples at a time with a final elution volume of 55 μl.

The method has been optimized to purify genomic DNA ready to use in PCR and fragment analysis reactions. Genomic DNA was purified on a Biomek® FX (Beckman Coulter, Fullerton, CA, USA) dual bridge. Hardware needed on this platform includes a gripper tool and a vacuum manifold. Steps such as binding condition set-up, wash and vacuum times have been optimized to deliver consistent results without contamination. Used lab wares were removed from the deck of the instrument by an ORCA® robotic arm (Beckman Coulter). The collar and manifold were adapted to allow elution into a standard microtiter® plate (Corning Incorporated). The Spacer Collar was used to stack the Plant Binding plate, the safe-spacer plate and, finally, the recovery plate.

2.3.2.2. DNA automated Whole Genome Amplification (WGA)

The following protocol, based on the OmniPlex™ WGA2 kit (Sigma), is the result of collaboration with the SZN Molecular Biology Service. The whole genome amplification (WGA) process was divided into fragmentation, OmniPlex™ library generation, and PCR amplification. The first two steps, fragmentation and library generation, should be carried out uninterrupted, as the ends of the library DNA can degrade quickly thus affecting the efficiency of subsequent steps. Experiments were performed along with a positive control DNA sample, such as the Control Human Genomic DNA included in the kit. A robotic procedure was developed using the BFX robotic platform and all pipetting steps were optimized. The reagent used for the preparation was 2-fold diluted fragmentation buffer (concentrated 10 times), made by

adding H_20, to create a stabilization solution, then mixing 1:1 with the library preparation, and then 1:1 with the enzyme solution. The plate was then setup for the amplification master mix.

Step 1 (fragmentation). A volume of 2 µl of 5x Fragmentation Buffer was added to 8 µl of DNA sample in a PCR 96-well plate (Unithermal full skirted 96 x 0.2 ml, Delchimica Scientific Glassware, code 20-5010310, Naples, Italy), mixed by pipetting four times. The plate was then placed in a thermal cycler at 95 °C for 4 min. Incubation time is very important, since small variations can affect the final results. After incubation, the plate was immediately cooled on ice and centrifuged briefly to collect the contents.

Step 2 (library generation). A volume of 2 µl of Library Preparation and stabilization solution mixture was added to each sample and mixed pipetting four times. Samples were spun down by centrifugation, and placed in a thermal cycler at 95 °C for 2 min before a first cooling on ice, followed by a short centrifugation and a further cooling. A 2 µl of Library Preparation Enzyme and buffer mixture were then added, and mixed by pipetting four times and briefly centrifuged. Samples were placed in a thermal cycler and incubated as follows: 16 °C for 20 min, 24 °C for 20 min, 37 °C for 20 min, 75 °C for 5 min, followed by 4 °C on hold. Samples were removed from the thermal cycler, centrifuged briefly and either immediately amplified or stored at -20 °C. OmniPlex™ library DNA, generated in stepped isothermal reactions, can be stored up to three days at -20 °C.

Step 3 (PCR amplification). Samples were thoroughly vortexed, briefly centrifuged, and subject to an initial denaturation at 95 °C for 3 min followed by a 14 cycles of denaturation at 94 °C for 15 seconds and Annealing/Extending at 65 °C for 5 min.

Completed reactions were momentarily stored at 4 °C prior to analysis or purification or stored at -20 °C.

2.3.2.3. DNA automated purification in 96-well plates

DNA purifications from plates were developed and performed at the SZN Molecular Biology Service on Multiscreen (HTS96) PCR filtration plate (Millipore). A robotic procedure was created to realize the experiment on the Biomek® FX robotic platform and all pipetting steps were optimized.

The filtration plate was used to purify WGA DNA-plate and PCR-plate reactions (*i.e.* from salts, primer excess, reactions secondary products, *etc.*). PCR reaction was loaded and filtered with a Millipore vacuum manifold for 5–10 min or until wells were dry. Each well was then filled with water (or buffer), the plate agitated by shaking and retrieved samples were purified by aspiration. The purification method was processed on a 96-well plate of PCR reactions in less than 20 minutes with a final recovery volume of 25 µl that can be directly used in downstream applications.

The method has been optimized to recover ~75% of the total loaded DNA. Timing and duration of steps such as washing, re-suspension and vacuum filtration have been optimized to deliver consistent results. MultiScreen PCR 96 filter plates (Millipore) provided high recovery for even the smallest miniaturized reactions. The plates used a size-exclusion membrane and vacuum filtration to provide a one-step protocol for excellent results. No centrifugation or precipitation steps were required.

2.4. Quantification of the DNA extracted and DNA plate

Nanodrop analysis was used to assess DNA purity whereas DNA quantification and PCR product specificity were assessed by agarose gel electrophoresis.

2.4.1. Nanodrop spectrophotometer estimations

A Nanodrop (Spectophotometer ND-1000, NanoDrop®, Wilmington, DE, USA) was used to assess the purity of DNA samples from the A_{260}/A_{280} ratio (the absorption of DNA (260 nm) compared to the absorption of RNA (280 nm)): ratios higher than 1.8 indicate that the DNA sample is of good quality (Fig. 2.2). The A_{260}/A_{230} ratio relates DNA (260 nm) to proteins (absorption at 230 nm): if the ratio is higher than 1.8, the sample is considered to be without major contaminants (Fig. 2.2).

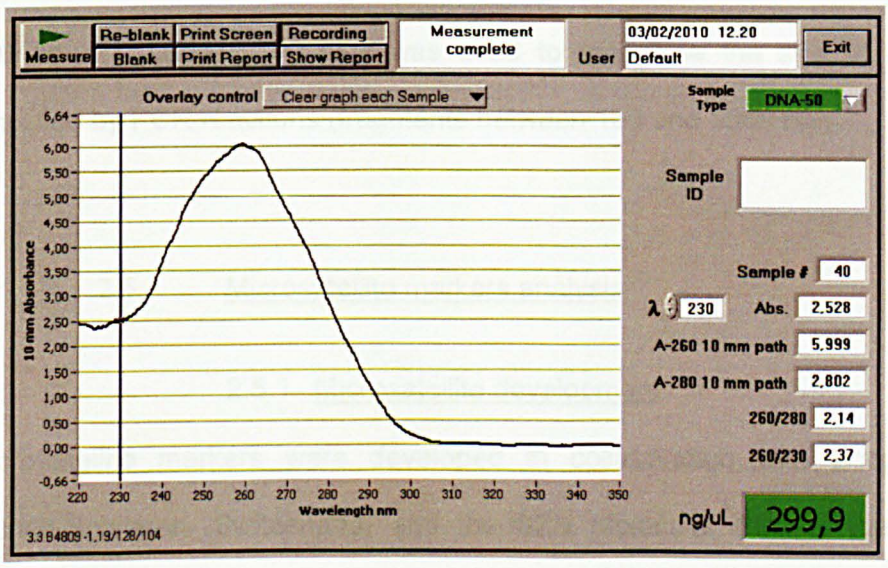


Figure 2.2: Nanodrop representation (sample E1588) for DNA quantification ($ng.\mu l^{-1}$) and evaluation of DNA purity through DNA-RNA (260 nm / 280 nm) and DNA-protein (260 nm / 230 nm) absorbance ratio. Upper than 1.8 DNA is considered as pure.

2.4.2. Agarose gel electrophoresis

DNA quality and quantity were estimated also by agarose gel electrophoresis. A 1.5 g of agarose (AppliChem, Darmstadt, Germany) was added to 150 ml of 0.5x TBE (Tris-Borate-EDTA) and dissolved to 10 $mg \cdot ml^{-1}$ by heating. Ethidium bromide (0.2 $\mu g.ml^{-1}$ final concentration, Sigma) was added before pouring the gel into the cast. Solidified gel was put into the electrophoresis tank and submerged in 0.5x TBE buffer. A final volume of 6 μl (1 μl of 6x DYE, plus DNA and H_2O, to a final volume of 6 μl) was loaded into each well. An initial 32 mA current was applied followed by 58 mA once the samples had entered the gel.

DNA fragment length and quantity were estimated in comparison to 1 μL of a DNA molecular size marker. Two different markers were used: Lambda II (Fig. 2.3) and Ladder100plus (Fig. 2.4) (Fermentas UAB, Vilnus, Lithuania). Lambda II (or Lambda/*Hind* III) was used to quantify the genomic DNA after extraction (from 6 to 250 $ng \cdot \mu l^{-1}$). Ladder 100 plus was used to recognise the size of microsatellites amplified by PCR reactions (fragments between 100 and 1500 bp).

2.5. Microsatellite markers analysis

2.5.1. Microsatellite development

Microsatellite markers were developed in collaboration with Ecogenics GmbH (Zürich-Schlieren, Switzerland) and the SZN Molecular Biology Service (Naples, Italy). Ecogenics delivered an enriched microsatellite library prepared using *Bam*H1, digested genomic DNA affinity-selected for di- and trinucleotide repeats and ligated into pUC19. Microsatellite sets and their amplification protocols were developed in collaboration with the SZN Molecular Biology Service.

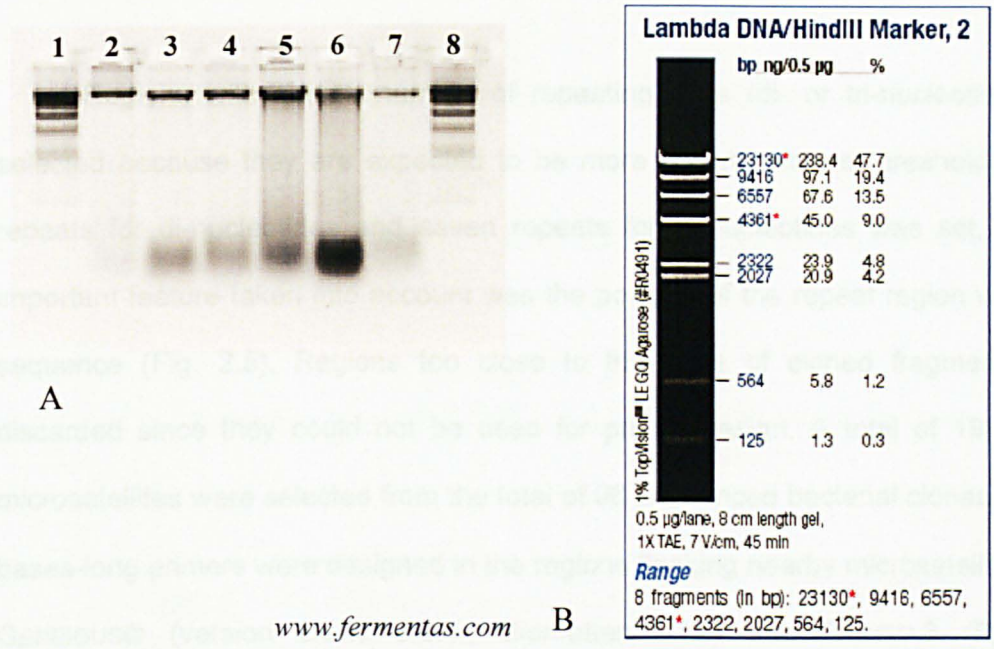


Figure 2.3: Quantification of DNA concentration: (A) by agarose gel electrophoresis (1st and 8th columns: Ladder Lambda II, 2nd to 7th columns: genomic DNA loaded after extraction (highest band: gDNA; lowest band: RNA residuals and/or DNA fragment), and (B) rule use (Ladder Hind III/Lambda II).

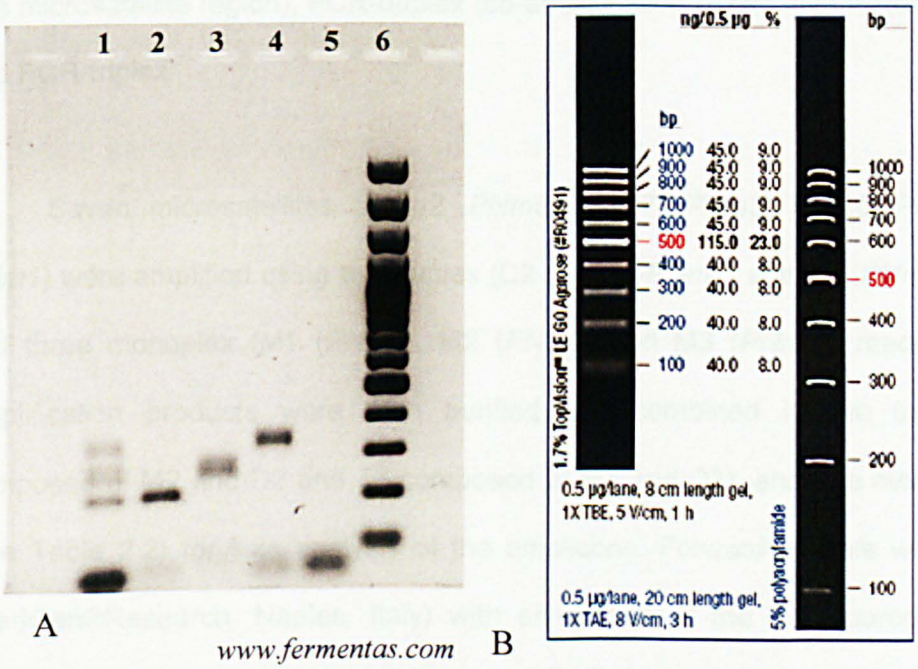


Figure 2.4: Evaluation of the fragment size of amplified gene fragments after PCR reaction: (A) by agarose gel (from the bottom to the top of each column: 1st column: dNTP, PNm3, PNm6 and PNm16, 2nd column: dNTP and PNm3, 3rd column: dNTP and PNm6, 4th column: dNTP and PNm16, 5th column: dNTP in reaction blank and 6th column: Ladder 100plus) and (B) rule use (Ladder 100plus: 100bp-1000bp).

Regions with a high number of repeating units (di- or tri-nucleotide) were selected because they are expected to be more polymorphic; a threshold of eight repeats for di-nucleotides and seven repeats for tri-nucleotides was set. Another important feature taken into account was the position of the repeat region within the sequence (Fig. 2.5). Regions too close to the ends of cloned fragments were discarded since they could not be used for primer design. A total of 19 putative microsatellites were selected from the total of 96 sequenced bacterial clones. Twenty bases-long primers were designed in the regions flanking nearby microsatellite, using GENEIOUS© (version Basic 3.6.1., Biomatters Ltd.) and PRIMER 3 (Rozen & Skalestsky, 2000), with the goal to generate PCR amplicons ranging between 300 and 400 bp. Microsatellite regions were amplified as PCR-monoplex (amplification of one microsatellite region), PCR-duplex (co-amplification of two microsatellite regions) and PCR-triplex.

Seven microsatellites (*PNm2, PNm5, PNm7, PNm3, PNm6, PNm16,* and *PNm1*) were amplified using two duplex (D2 (*PNm5-PNm7*) and D3 (*PNm3-PNm16*)) and three monoplex (M1 (*PNm1*), M2 (*PNm2*) and M3 (*PNm6*)) reactions. PCR-amplification products were then purified and combined in two triplexes (T2 composed of M2 and D2 and T3 composed of M3 and D3), and one monoplex (M1) (see Table 2.2) for size analysis of the amplicons. Forward primers were labelled (MeriGen®Research, Naples, Italy) with either one of the two fluorophores: Cy5 (blue) and IRD700 (green) (Fig. 2.6).

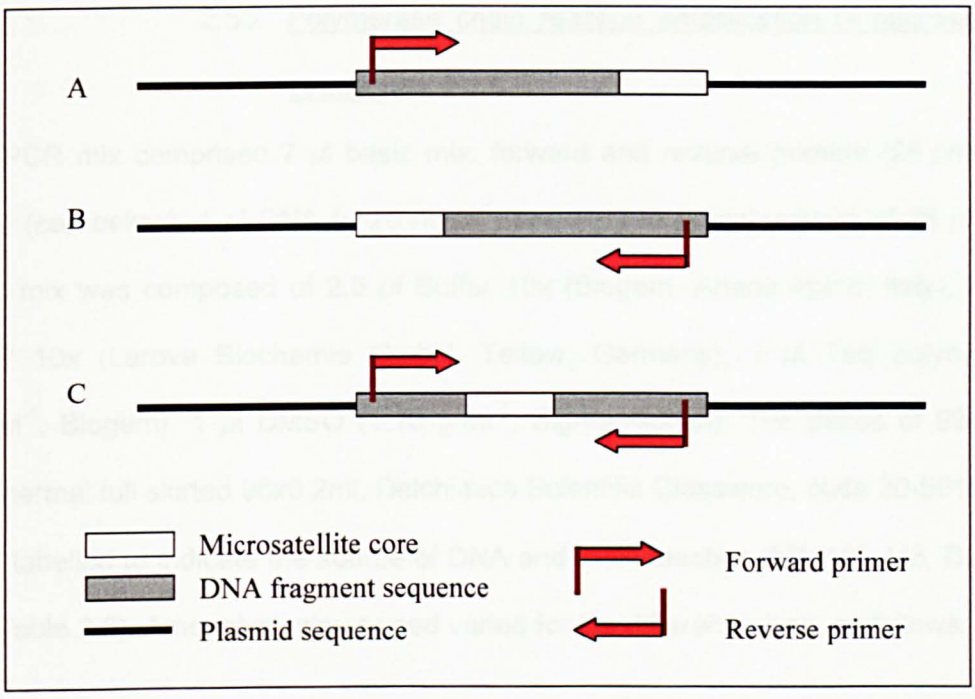


Figure 2.5: Cartoon that explains the possible configuration of microsatellites position after cloning. Black bold line indicates the plasmid sequence; gray box the DNA fragment sequenced and the white box the microsatellite core repeats; orange arrows indicate the primers forward and reverse. A and B are two configurations to be omitted because the amplification or sequencing of these microsatellite, too close to the primer sequence, may be not correctly sequenced. C is the best configuration.

Table 2.2: Microsatellites primers (Locus name (PNm), Forward primer sequence (F), Reverse primer sequence (R), microsatellites size and core repeats in base pair (bp), the fluorophore (Fluo), the melting temperature (T°m), the sequence (5'-3'), the annealing temperature (T°a), PCR (P) and Fragment analyser (F) combination (monoplex (M) - duplex (D) - triplex (T) combination).

Locus name	Primer (bp)		Size (bp)	Core	Fluo	T°m (°C)	Sequence 5' to 3'	T°a (°C)	P	F
PNm1	F	22	149	$(TC)_{23}$	IRD700	60.4	CACCAATTGCATCCTAAAAGGG	54	M1	M1
PNm1	R	24			None	59.4	TCCGTCTAAGCCTGTATTTGTGAC	54		
PNm2	F	21	189	$(AC)_{17}$	IRD700	60.7	GGGATCGATTCGTGAAAGAGC	58	M2	T2
PNm2	R	21			None	61.6	GCATAGAAGCACGGCACAGTG	58		
PNm3	F	23	208	$(GAC)_{8}$	Cy5	60.5	GGATCGAATAGGGGATGAATACG	58	D3	T3
PNm3	R	22			None	60.4	GGAGCTTGCATCATCATCACAG	58		
PNm5	F	21	240	$(GT)_{11}$	Cy5	60.8	GAACAGAACTGCCCGAAGGAC	58	D2	T2
PNm5	R	21			None	59.8	AGGATCACCCACGAGACACTG	58		
PNm6	F	22	251	$(CT)_{9}$	IRD700	60.5	AGCGAAAGCGACAAATAGCATC	58	M3	T3
PNm6	R	21			None	60.2	TGAGCAAAAGGACGAAACGAG	58		
PNm7	F	20	260	$(CA)_{8}$	IRD700	60.1	GTTGGCACCGGTGGTCTAAC	58	D2	T2
PNm7	R	19			None	60.0	CTTCGACGCTCCATTGGTG	58		
PNm16	F	23	336	$(GTC)_{7}$	Cy5	60.7	GGATCATACTGGAGGGGAACAAG	58	D3	T3
PNm16	R	24			None	60.0	GCTTTCACATCCAGAAGACAACAG	58		

2.5.2. Polymerase chain reaction amplification of microsatellite sequences

The PCR mix comprised 7 μl basic mix, forward and reverse primers (25 *p*mol·μl^{-1} each) (see below), 1 μl DNA (~ 20 ng·μl^{-1}) and H_2O to a final volume of 25 μl. The basic mix was composed of 2.5 μl Buffer 10x (Biogem, Ariano Irpino, Italy), 2.5 μl dNTP 10x (Larova Biochemie GmbH, Teltow, Germany), 1 μl Taq polymerase (3 u·μl^{-1}, Biogem), 1 μl DMSO (1.10 g·ml^{-1}, Sigma-Aldrich). The plates of 96-wells (Unithermal full skirted 96x0.2ml, Delchimica Scientific Glassware, code 20-5010310) were labelled to indicate the source of DNA and PCR-reaction (M1, M2, M3, D2, D3; see Table 2.2). Amount of primer used varied for the different mixes, as follows:

- M1, M2 and M3: 1 μl of forward (Fw) primer (25 *p*mol·μl^{-1}, IRD700) and 1 μl of reverse (Rev) primer (25 *p*mol·μl^{-1}).

- D2 (*PNm5-PNm7*): 1.5 μl of *PNm5* Fw (25 *p*mol·μl^{-1}, Cy5), 1.5 μl of *PNm5* Rev (25 *p*mol·μl^{-1}), 1 μl of *PNm7* Fw (25 *p*mol·μl^{-1}, IRD700) and 1 μl of *PNm7* Rev (25 *p*mol·μl^{-1}).

- D3 (*PNm3-PNm16*): 1 μl of *PNm3* Fw (25 *p*mol·μl^{-1}, Cy5), 1 μl of *PNm3* Rev (25 *p*mol·μl^{-1}), 3 μl of *PNm16* Fw (25 *p*mol·μl^{-1}, Cy5) and 3 μl of *PNm16* Rev (25 *p*mol·μl^{-1}).

PCR-amplification conditions were as follows:

- M2, M3, D2 and D3: 96 °C for 6 min, 30 cycles of denaturation-annealing-extension-elongation at respectively 94 °C for 1 min, 58 °C for 1 min and 72 °C for 2 min, followed by a final extension step at 72 °C for 7 min, and 15 min at 25 °C then overnight at 4 °C.

- M1: 96 °C for 6 min, 40 cycles of denaturation-annealing-extension-elongation at respectively 94 °C for 1 min, 54 °C for 1 min and 72 °C for 2 min, a step at 72 °C for 7 min, followed by 15 min at 25 °C and overnight at 4 °C.

The PCR products were purified (see section 2.3.2.3) prior to fragment analysis. Purification was achieved using the Millipore Multiscreen HTS PCR 96-well plate kit (Millipore) and a robotic station (Beckman Coulters Biomek® FX Laboratory Automation Workstation) equipped with ORCA® robotic arm (Beckman Coulter, Fullerton, CA).

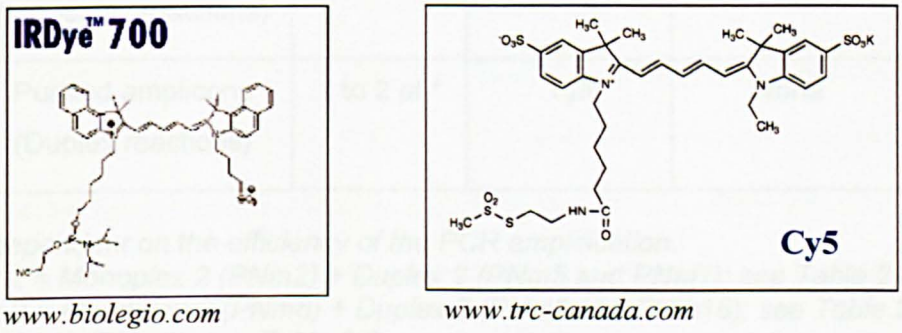


Figure 2.6: Fluorochrome molecule for fragment analysis (IRDye700 (green) and Cy5 (blue) fluorochrome).

2.5.3. Fragment analysis of amplified microsatellite loci.

Purified PCR products were combined into two triplexes (T2 and T3) and one monoplex (M1, *PNm*1) (Table 2.2). A robotic protocol for the BFX robotic platform and all pipetting steps were optimized. The fragment analyser mix comprised a basic mix (36.8 µl Formamide and 0.2 µl Size standard (GenomeLab™ DNA size standard Kit-400, Beckman Coulter®)) and amounts of PCR-product adjusted as follows (Table 2.3): M1: 2.5 µl of M1, T2: 1-3 µl of M2 and 1-2 µl of D2, and T3: 2 µl of M3 and 1 µl of D3. The fragments were analysed on a Beckman Coulter CEQ™2000XL DNA Analyzer (denaturation 90 °C 120 sec, Injection 2 Kv 60 sec, Voltage 6 Kv 80 min).

Table 2.3: Amplified microsatellite sample preparation for fragment analysis (CEQ2000XL DNA Analyser). Mix is composed of Formamide, Size standard and PCR-reactions purified.

	Triplex 2 §	Triplex 3 §§	Monoplex 1 §§§
Formamide	~ 36.8 µl	~ 36.8 µl	~ 36.8 µl
Size standard	0.2 µl	0.2 µl	0.2 µl
Purified amplicons (Monoplex reactions)	1 to 3 µl *	2 µl	2.5 µl
Purified amplicons (Duplex reactions)	1 to 2 µl *	1 µl	None

** range dependent on the efficiency of the PCR amplification.*
§ Triplex 2 = Monoplex 2 (PNm2) + Duplex 2 (PNm5 and PNm7); see Table 2.2
§§ Triplex 3 = Monoplex 3 (PNm6) + Duplex 3 (PNm3 and PNm16); see Table 2.2
§§§ Monoplex 1 (PNm1); see Table 2.2.

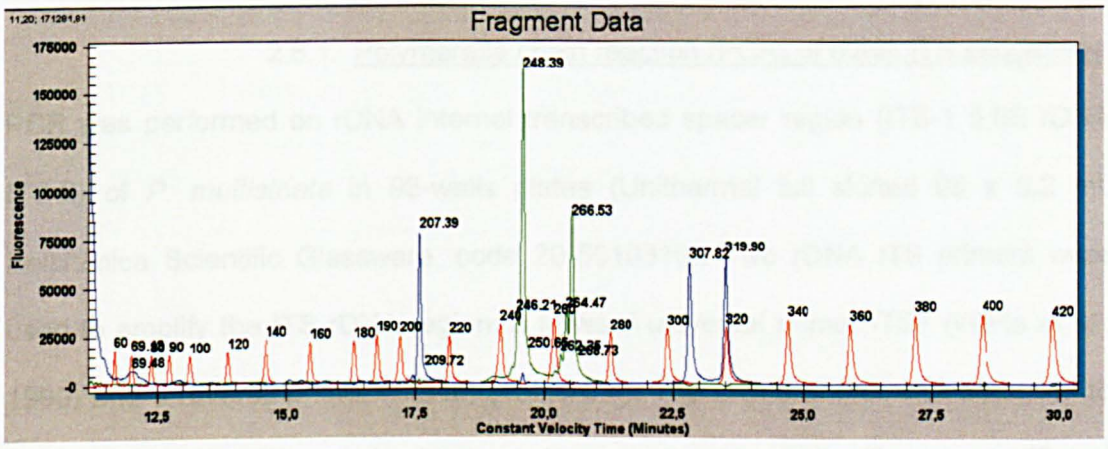


Figure 2.7: Example of triplex fractionation to resolve microsatellite amplicons: PNm3 (blue), homozygote with 207 bp fragment; PNm6 (green), heterozygote with 249 and 267 bp fragments; PNm16 (blue), heterozygote with 307 and 320 bp fragments. Size standard (red).

Fragment size analysis was guided by the size standard rule (0-400 bp, red colour). Microsatellite amplicon quantity and migration were visualised using the fluorophor incorporated into the Forward primers (Cy5 blue colour and IRD700 green colour (Fig. 2.6)). Homozygous microsatellite loci produce a single peak whereas heterozygous loci produce two distinct peaks (Fig. 2.7).

CEQ™2000XL (version 4.3.9, Beckman Coulter™) was used to infer microsatellite length, LOSITAN (Beaumont & Nichols, 1996) to test microsatellite neutrality, GIMLET (version 1.3.3, (Valière, 2002)) to assess genotypic diversity, GENALEX (version 6, (Peakall & Smouse, 2006)) to estimate population genetic parameters, GENETIX (version 4.05.2) to calculate linkage dis-equilibrium, MICROCHECKER (version 2.2.3, (Shipley, 2003)) to identify null alleles and STRUCTURE (version 2.1, (Pritchard *et al.*, 2000)) to derive population structures (see Chapter 6).

2.6. Internal Transcribed Spacer (ITS) region rDNA analysis

2.6.1. Polymerase chain reaction (PCR) of rDNA ITS sequences

PCR was performed on rDNA internal transcribed spacer region (ITS-1 5.8S rDNA ITS-2) of *P. multistriata* in 96-wells plates (Unithermal full skirted 96 x 0.2 ml, Delchimica Scientific Glassware, code 20-5010310). Two rDNA ITS primers were used to amplify the ITS rDNA region: a forward universal primer ITS1 (White *et al.*, 1990) and a reverse primer (ITS4b) (Table 2.4). The ITS4b primer was designed to optimise the reaction and obtain a larger amount of amplicons specific of *P. multistriata*. The design of ITS4b took into the account the initial position of the universal reverse primer ITS4 (White *et al.*, 1990) on the rDNA gene and

P. multistriata sequences (D'Alelio *et al.*, 2009a). The primer design was performed using BioEdit (version 7.0.9.0, (Hall, 1999)) and Primer 3 (Rozen & Skalestsky, 2000). Amplification reactions were performed in a 25 μl final volume comprising 14 μl sterile water, 2.5 μl Buffer (10x), 2.5 μl dNTP (10x), 2 μl Taq polymerase (3 $u \cdot \mu l^{-1}$, Biogem), 1μl of each primer (25 $pmol \cdot \mu l^{-1}$), 1 μl DMSO (1.10 $g \cdot ml^{-1}$), and 1 μl DNA (25 $ng \cdot \mu l^{-1}$). PCR included a denaturation step at 96 °C for 2 min, 40 cycles of denaturation-annealing-extension at 94 °C for 1 min, 46 °C for 1 min and 72 °C for 2 min, a final extension step at 72 °C for 5 min followed by 15 min at 25 °C and 4 °C overnight. The PCR products were purified before being sequenced. Purification was performed using the Millipore Multiscreen HTS PCR 96-well plate kit (Millipore) and a robotic station (Beckman Coulters Biomek® FX Laboratory Automation Workstation) equipped with ORCA® robotic arm (Beckman Coulter).

Table 2.4: rDNA ITS primers sequence (5' – 3'): ITS1 as universal forward primer and ITS4b as designed reverse primer for the amplification of ITS-region of Pseudo-nitzschia multistriata species.

Primer	Reference	Sequence 5' to 3'	Bp
ITS1	White *et al.*,1990	TCCGTAGGTGAACCTGCGG	19
ITS4b	This study	TCCTCCGCTTAATTATATGC	20

2.6.2. Fragment analysis of rDNA ITS sequences

The rDNA ITS sequences were obtained using a BigDye Terminator Cycle Sequencing technology (Applied Biosystems, Foster City, CA). Sequencing products were purified using the Agencount CleanSEQ Dye terminator removal kit (Agencourt Bioscience Corporation, 500 Cummins Center, Suite 2450, Beverly, MA, USA) and

the robotic station Biomek® FX (Beckman Coulter). Products were analysed using an automatic capillary electrophoresis sequencer 3730 DNA analyser (Applied Biosystems, Foster city, USA). BioEdit (version 7.0.9.0, (Hall, 1999)) was used to infer ITS electropherograms sequence.

CHAPTER THREE

OPTIMISATION OF CULTURE MAINTENANCE

3.1. Introduction

Diatoms constitute study objects for a range of research lines and have great potential for technical development. Their composite silica cell wall, called a frustule, is the subject of material studies, as industry is still unable to construct microscopically small silica elements under ambient conditions and in the exquisite details diatoms do (Hildebrand, 2008; Kröger & Poulsen, 2008). Furthermore, diatoms can grow extremely fast, and are therefore prime objects for micro-algal mass culture tooled towards the production of high quality substances (silica compounds, toxins, *etc.*). A prerequisite for all these research efforts is the access to strains over extended periods (years) to revisit, or build upon, results of earlier studies conducted on the same strains. Strains can become models of their taxa on which interdisciplinary research is conducted over extended periods.

The silica frustule consists of two parts, called thecae that fit precisely into one another: one called a hypotheca and the other, larger, called an epitheca, the whole resembling a Caprice-des-Dieux® cheese-box. Each theca consists in its turn of a valve and an accompanying set of girdle bands. Although each of these elements are rigid, they permit cell growth because the elements can slide alongside one another, and the girdle bands can be constructed and deposited in the cell wall as the cell grows. The pair of theca permits cell division whilst the living cell remains completely surrounded by the frustule elements even during cell division. Upon division, each of the daughter cells inherits one of the thecae from the mother cell and constructs a

new theca, fitting precisely inside it (MacDonald, 1869; Pfitzer, 1869; Round *et al.*, 1990). Thus, the daughter cell that has inherited the mother cell's epitheca will be of the same size as the mother cell whereas the one that inherited the hypotheca will be a tiny bit smaller. With ongoing vegetative growth, diatom cells in a culture will thus become smaller and the variance around their average cell size will increase. The progressive miniaturisation is referred to as the MacDonald-Pfitzer rule of diatom growth (MacDonald 1869; Pfitzer 1869). The principal way to escape from becoming too small is to enter in sexual reproduction because the zygotes swell up to re-establish initial cell size (Chepurnov *et al.*, 2004). Only a few diatom lineages can escape from this miniaturization. Some *Thalassiosira* species can shed their cell wall during the vegetative cycle and then swell up like a zygote and form a new cell wall (von Dassow *et al.*, 2006) whereas *Phaeodactylum* lacks a rigid cell wall, and therefore, escapes from miniaturization (Lewin *et al.*, 1958). Strains of these species, thanks to cell enlargement, can be maintained indefinitely as clones, and it is not surprising that they have become model species for multidisciplinary studies (Chepurnov *et al.*, 2008).

The miniaturization constitutes a severe setback to the culture of diatoms; strains of most species cannot be maintained as clones indefinitely. Rapid growth of a monoclonal culture implies rapid miniaturization of the cells and hence a reduced lifetime of the monoclonal strain. To prolong its lifetime a strain has to be maintained under slow growth conditions whilst ensuring survival. Such a way of strain maintenance has the practical advantage of less frequent culture transfer, and hence time and expenses saved on maintenance.

Strain maintenance of *Pseudo-nitzschia* species is no sinecure, and strains do not keep for long because of the miniaturization. For these reasons they are rare in

culture collections in spite of the ecological and societal importance of the genus and the ongoing research interest in it. Vegetative enlargement has never been observed in *Pseudo-nitzschia* and therefore, the only way to maintain monoclonal research strains of *Pseudo-nitzschia multistriata,* the species in the focus of this Ph.D. thesis over a long period, is to minimize the growth rate whilst ensuring survival. A few hundred strains cultures were established from environmental samples and maintained over a period of several months. The DNA was harvested for PCR amplification of the ITS rDNA marker regions (Chapter 4) and microsatellite markers (Chapters 5 and 6). Prolonged maintenance was needed to permit re-extraction of DNA if necessary, and to permit follow-up experiments after the first results had been evaluated.

The two principal categories of factors believed to affect reliable long-term maintenance of *P. multistriata* cultures are those that ensure their wellbeing in the natural environment (Suzuki & Johnson, 2001) and those that permit only slow growth without reducing their chances of survival. Environmental conditions during the blooming seasons and during the winter period at the Long Term Ecological Research site MareChiara (LTER-MC) provide clues into what these long-term maintenance conditions are for *P. multistriata*. At the LTER-MC site, the species is observed usually in early summer (June-July) and autumn (September-November) during which the water temperature is at respectively *ca.* 21-24 °C and 17-24 °C and the photoperiod 14:10 (L:D; in early-summer) and 12:12 to 9:15 (through the autumn-bloom; Fig. 3.1). Since *Pseudo-nitzschia* species are not known to form resting stages, the population has to survive 14-15 °C, the local water temperature during midwinter. This is also the water temperature below the thermocline, where the species might survive during the high summer.

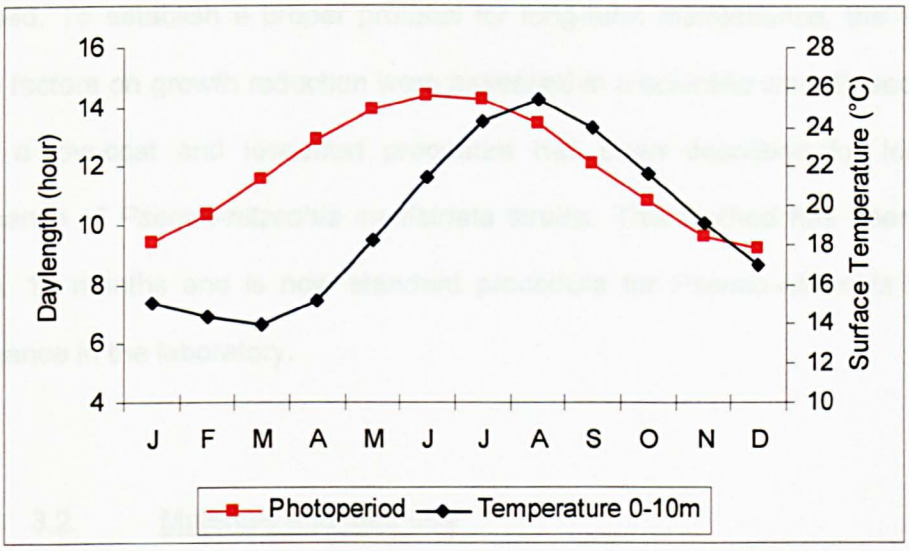


Figure 3.1: Average monthly day-length (red line) and temperature (blue line, 0 to 10 m) over the annual cycle at the LTER-MC (1984-1990, EEP-data).

The wellbeing of *P. multistriata* were assumed to be guaranteed when irradiance, day-length, temperature and turbulence levels approach those in the water column during the period in which the species is recorded in the plankton and when nutrients are present in sufficient concentrations. Lowering the temperature and light intensity are the principal options to reduce the growth rate (D'Alelio et al., 2009b). It should be taken into the account that culture conditions differ radically from those in the field; nutrient concentrations in F/2 medium are much higher, precipitates often form in autoclaved seawater, grazers and protistan antagonists are absent, and bacteria, if given a chance, may swamp the diatoms in a culture and affect their survival (Obernosterer & Herndl, 1995).

A total of 535 strains of *P. multistriata* were sampled over the bloom periods of 2008 and 2009, and the cultures of all these strains initially needed to be transferred once a week. To slow down growth, the temperature and irradiance levels were

decreased. To establish a proper protocol for long-term maintenance, the effect of multiple factors on growth reduction were assessed in a scientific way. Based on the results, a low-cost and low-effort procedure has been described for long-term maintenance of *Pseudo-nitzschia multistriata* strains. This method has been tested over *ca.* 15 months and is now standard procedure for *Pseudo-nitzschia* species maintenance in the laboratory.

3.2. Materials and Methods

3.2.1. Culture establishment

Cultures of *Pseudo-nitzschia multistriata* were generated from single chains of cells or from single cells in net-samples (Chapter 2). Cells or chains were washed by means of serial transfer through a series of sterile seawater droplets using sterile drawn-out Pasteur pipettes, an inverted binocular microscope (100x magnification; Leica Microsystems DMIL, Wetzlar, Germany) and an hourglass slide. Single cells or chains were then transferred to 2 ml sterile filtered F/2 medium in 12-well culture plates (Corning® Costar® CLS3513, NY, USA) and incubated under ambient shadowed light at 20 °C on a North-facing windowsill (*i.e.*, no direct sunlight). As soon as possible, strains were screened for growth and unialgal status. Unialgal strains were transferred to – and further maintained in – 25 ml bottles (Corning® Flask, 25 cm^2, code 430639, Corning Inc., NY, USA) filled with F/2 medium.

3.2.2. Medium preparation

F/2 medium was made from oligotrophic Mediterranean seawater collected offshore in the Gulf of Naples. The water was filtered over a 0.45 μm filter (Millipore S.p.A.,

nitrocellulose membrane code HAWP0900, Milan, Italy) and then autoclaved. Salinity was adjusted to 36 psu by adding sterile milli-Q water, and F/2 was obtained through addition of aliquots of 20 ml of 50x concentrated F/2 (Table 2.1, Sigma Aldrich S.r.l.) per litre. The F/2 medium was filtered over a 0.2-μm filter (Millipore, Filter Stericup-GP SCGPU05RE, Billerica, MA, USA) just before use in order to eliminate precipitates.

3.2.3. Parameters set up

The factors tested were: a) temperature, b) irradiance, c) mode of agitation and d) inoculum size. Quantitative results were obtained for the temperature and inoculum size parameters. The other parameters were evaluated qualitatively. If growth was observed in 25 ml bottles under the experimental conditions, then the experiment was repeated in 12-well plates.

3.2.3.1. Temperature

The temperatures tested were 10 °C (far below the minimum temperature of the water column at the LTER-MC during winter), 15 °C (the minimum winter temperature, and temperature under the thermocline in summer) and 22 °C (the average temperature above the thermocline during blooming periods) (Fig. 3.1, Table 3.1). In order to estimate the growth rate of *P. multistriata* strains, four cultures (SY413, SY416, SY486 and SY487) were transferred in bottles filled with 200 ml of F/2 medium (Corning® Flask, code 430823, Corning Inc., NY, USA) with an inoculum of *ca.* 1300 cells·ml^{-1}, and placed at a photoperiod of 12:12 (L:D) under a sinusoidal light regime comprising an impulse from dark to light (40 μmol photons·m^{-2}·s^{-1}) and

12 hours of light with a maximum irradiance at 80 μmol photons·m^{-2}·s^{-1} after 6 hours. The bottles were agitated gently twice a week to reduce cells from accumulating on the bottom of the bottle. Growth was evaluated by observation and comparison, and if required, by cell counts from fixed samples (three millilitres of culture were fixed with 60 μl acid formaldehyde). Cell counts were performed using a 1-ml Sedgewick-Rafter chamber and counted using an Axioskop2 plus microscope (Zeiss, Göttingen, Germany) under 10x objective magnification. The growth curve (Fig. 3.2) was assessed as follows: the exponential growth period was estimated from a semi-log plot of the cell concentration (cells·ml^{-1}) over time. A least-square regression was applied to the selected log (base 10) data. A logarithmic transformation from log base 10 to log base 2 was used to estimate the number of divisions per day (Guillard, 1973).

3.2.3.2. Light regime

Two light regimes, sinusoidal and on/off, have been tested over culture maintenance at 15 °C. Maximum irradiances on on/off light regime were chosen at 30, 40, 50-60, 80 and 100 μmol photons·m^{-2}·s^{-1}. A sinusoidal light regime was chosen to mimic the irradiance rhythm for a photoperiod of 12:12 (L:D) during the autumnal *P. multistriata* bloom. Experiments were conducted in 25 ml bottles (Corning® Flask, 25 cm^2, code 430639, Corning Inc., NY, USA) and qualitatively tested.

3.2.3.3. Mode of agitation

The agitation rate was tested on well-growing cultures in 12-well plates (Corning® Costar® CLS3513, NY, USA) and placed at 15 °C, 12:12 L:D and a sinusoidal light

regime with a maximum of 80 μmol photons·m^{-2}·s^{-1}; light intensity at which slow but normal growth was observed and cells appeared healthy. Three agitation regimes were tested to assist the cells to remain suspended and to avoid accumulation of bacteria (i) reciprocating agitation (Heidolph-Instrument, Promax 1021, Milan, Italy), (ii) orbital agitation (S01 Stuart and SSL1; both from International PBI S.p.A., Milan, Italy), both at 30, 50, 70 and 80 rpm and (iii) no agitation at all. The agitators were selected because of their adjustable velocity and their propensity to not generate heat when working.

3.2.3.4. Inoculum size

The optimal number of cells in the inoculum (*i.e.*, the volume of the inoculum) was explored using 12-well plates placed at 15 °C, 12:12 L:D under a sinusoidal light regime with a maximum irradiance of 80 μmol photons·m^{-2}·s^{-1} and orbital agitation at 70 rpm. The number of cells in the inoculum was decreased until the daughter cultures, resulting from the inoculated cells, grew for *ca.* one month. The condition of the cultures was monitored weekly, using a light microscope (100x magnification; Leica Microsystems DMIL, Wetzlar, Germany). The mean number of cells in the inoculum (Table 3.2) was calculated by multiplying the inoculated volume with the cell concentration observed in the mother-culture.

3.2.4. Culture maintenance (Fig. 3.3 to 3.6)

Culture maintenance protocol was performed on *P. multistriata* from initial cells ($N_{STRAINS}$ = 68 strains), *P. multistriata* smaller cells ($N_{STRAINS\ TOTAL}$ = 535 strains), and other diatoms species ($N_{STRAINS\ OTHER\ SPECIES}$ = 41 strains, *i.e. P. delicatissima* (clades 1 and 4), *P. pseudodelicatissima, P. mannii, P. dolorosa, P. calliantha, and*

P. galaxiae) and *Chaetoceros danicus*) respectively over 5, 15.5 and 4 months and the efficiency of the procedure was calculated (Tables 3.3 to 3.5).

Cultures were maintained in 12-well incubation plates. The transfer procedure was performed in semi-sterile conditions on a cleaned lab bench in a laminar flow hood. Wells were rinsed each with 2 ml of filtered F/2 medium for a few minutes to eliminate any contaminants on the plastic surface, upon which the rinsing medium was discarded. Wells were then filled with 2 ml medium using a 25 ml-pipette (Costar ® Stripette ®, code 4489, Corning Inc).

3.2.4.1. Stock maintenance (Fig. 3.4)

A stock culture ideally consists of medium with cells in exponential growth or late exponential growth. For each culture transfer, micro-volume pipette tips (0.1-10 µl, Corning Inc., code 4894) and a 0-2 µl Gilson pipette (Gilson Inc., Middleton, WI, USA) were used. Prior to transfer, the stock culture was mixed thoroughly by stirring the culture with a Gilson pipette tip. Then the pipette tip was drained and a small aliquot (1.2 µl) was taken from the mix and transferred to a new well with fresh medium. Aliquots were taken while pulling the pipette tip from the bottom to the surface of the stock culture. An aliquot of about 76 cells was transferred in 1.2 µl for categories 4 and 5 and 2.4 µl for categories 1 to 3 (see Table 3.2 for explanation of categories). The new culture plates were closed hermetically with Parafilm (Menasha, WI, USA), placed in piles in a climate-controlled growth chamber under experimental conditions. Day-length (12:12, L:D) and irradiance from the sides (neon, cool white light, Philips T94, Eindhoven, the Netherlands) were kept constant throughout the maintenance procedure.

3.2.4.2. Back-up maintenance (Fig. 3.5)

At each transfer of the cultures, a back-up plate was established as follows: after removing a small volume used to inoculate the new culture plate, the well content of the stock-plate was mixed once more and then emptied using plastic Pasteur pipettes (LP, code 137038). In this way, only a thin layer of medium with a small number of cells remains in the well. The well was refilled with 2 ml of fresh medium as soon as possible and incubated as illustrated above.

3.3. Results

3.3.1. Parameter set up

3.3.1.1. Temperature

Table 3.1: Growth rates of strains SY413, SY416, SY486 and SY487 measured at three different incubation temperatures (22 °C, orange; 15 °C, green; 10 °C, blue)). Step denotes the length of the exponential growth phase (EP, in days), the division rate (in divisions per day), the $S_{INITIAL}$ and $S_{PLATEAU}$ indicate the average frustule length (in µm) at the beginning of the exponential phase (EP) and at the stationary phase, respectively.

Temp. (°C)	Strain code	EP (days)	Division rate (divisions·day^{-1})	$S_{INITIAL}$ (µm)	$S_{PLATEAU}$ (µm)
22	SY486	7	0.69	40.3 ± 0.24	39.4 ± 0.33
22	SY487	7	0.70	44.0 ± 0.23	44.4 ± 0.55
Mean		7	0.70	42.2 ± 0.00	41.9 ± 0.78
15	SY413	11	0.37	40.8 ± 0.34	40.4 ± 0.31
15	SY416	7	0.42	40.6 ± 0.29	40.2 ± 0.39
15	SY486	7	0.61	40.3 ± 0.24	39.3 ± 0.37
15	SY487	7	0.48	44.0 ± 0.23	43.2 ± 0.41
Mean		8	0.47	41.4 ± 0.47	40.8 ± 0.52
10	SY413	0	0	40.8 ± 0.34	
10	SY416	0	0	40.6 ± 0.29	
10	SY486	0	0	40.3 ± 0.24	

For all the strains tested, death within a week was observed at 10 °C, a low division rate at 15 °C and a higher division rate at 22 °C (Table 3.1). Rates were independent of cell length because if length affected division rate, then SY413, SY416 and SY486 should have exhibited similar growth rates (Table 3.1, Fig. 3.2). Strains tested at 22 °C showed markedly higher division rates. Strains reacted differently during the maintenance culture; for example SY416 showed temporally interrupted growth in one experiment with two exponential phases (Fig. 3.2); a division rate of 0.42 over the first seven days, a complete arrest for 11 days, and a division rate of 0.38 over the final nine days of the experiment. Culture vitality was monitored at 15 °C up to 33 days (Fig. 3.2) for the four strains tested. Based on these results, the temperature condition for slow growth was chosen at 15 °C and this temperature was used in the follow-up experiments to select the remaining parameters.

3.3.1.2. Light regime

The on-off light regime resulted in erratic and unpredictable growth while the sinusoidal rhythm resulted in predictable low growth, uniform distribution and healthy cells when incubated at a maximum irradiance level of 80 μmol photons·m^{-2}·s^{-1} and for a photoperiod of 12:12 L:D. Cultures maintained at a maximum irradiance of 30 and 100 μmol photons·m^{-2}·s^{-1}, on/off light regime, invariably showed condensed cytoplasm and empty cells, indicating moribund or dead cells. Shadowed cultures grew markedly slower than cultures directly facing the light source. At a maximum irradiance of 40 and 80 μmol photons·m^{-2}·s^{-1}, on/off light regime, cells grew more slowly and markedly faster, respectively, than at 50-60 μmol photons·m^{-2}·s^{-1} but the cells appeared less healthy. Thus, a sinusoidal light regime settled with a dark at

0 μmol photons·m^{-2}·s^{-1}, an impulse from the dark to light to 40 μmol photons·m^{-2}·s^{-1}, with a maximum irradiance level set at 80 μmol photons·m^{-2}·s^{-1} and a photoperiod of 12:12 L:D, was selected as the optimal condition for slow but steady growth and this condition was used in the follow-up experiments to select the remaining parameters.

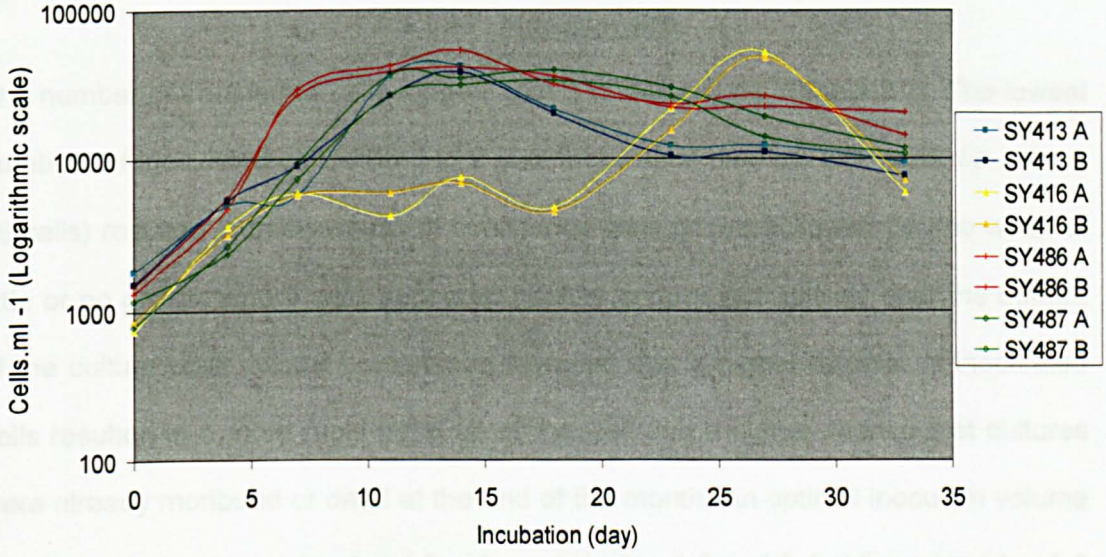


Figure 3.2: Growth curves of four strains in duplicate (A and B) at 15 °C (SY413 (light/dark blue, rectangles), SY416 (yellow/gold, triangles), SY486 (orange/red, plus) and SY487 (light/dark green, diamonds)). The vertical axis indicates the number of cell per ml and the horizontal axis indicates the incubation time in culture (in days).

3.3.1.3. Mode of agitation

If cultures in wells and bottles were not agitated, cells accumulated on the bottom became entrapped in mucilage and perished. Bacteria tended to proliferate in the unmixed medium, affecting survival of the diatoms. Under orbital agitation at 60 and 70 rpm cells remained thoroughly mixed and showed continuous growth. Instead, at 30 and 50 rpm, cells tended to accumulate in the centre of wells. Reciprocating agitation resulted in a less uniform distribution of the cells in the wells at 30 and

50 rpm than observed using orbital agitation and spill-over among wells was observed at 80 rpm. Thus, continuous orbital agitation at 70 rpm was selected as the condition to be used in the follow-up experiments to select the remaining parameters.

3.3.1.4. Inoculum size

The number of transferred cells ranged from 5 to 220 cell.ml^{-1} (Table 3.2). The lowest numbers of inoculated cells tested (1.2 μl of inoculation medium with cells; *i.e.*, about 76 cells) resulted in three weeks of continuous slow growth followed by one week of little or no growth where cells appeared healthy and moved actively over the bottom of the culture plate. Visual comparison revealed that a higher number of inoculated cells resulted in a more rapid filling up of the well and a higher chance that cultures were already moribund or dead at the end of the month. An optimal inoculum volume for slow growth was retained at 1.2 μl for categories 4-5 and 2.4 μl for categories 1-3 (Table 3.2). These categories of cell densities were established to allow rapid estimation of the inoculum volume needed through visual estimation of the cell densities. Note that any procedure to count the cells in the inoculum would be too time-consuming and thus impractical.

Table 3.2: Estimated cell concentrations by visual observation for cultures that need to be transferred on a monthly basis. The first column indicates the category (from 0 to 5) of cell concentration as estimated by visual examination (second column). In column three is reported the number of strains counted to estimate the average cell concentration (column four).

Category	Visual examination	Number strains counted	Mean number cells·ml^{-1}
0	Absence of cells or dead cells	0	0
1	Cells grouped in the centre	3	1 956
2	Cells covering almost all the bottom of the well	3	6 050
3	Cells covering all the bottom of the well	3	12 333
4	Cells covering the bottom of the well as well as suspended in the water	3	111 378
5	Very high concentrations of cells at the bottom of the well as well as suspended in the water	4	186 183

3.3.2. Culture maintenance

The following culture protocol was found to permit reliable long-term maintenance of *P. multistriata* strains in 12-well plates: transfer strains every four weeks, inoculation of a low number of cells (*ca.* 76) to avoid rapid crowding, growth under continuous orbital agitation (70 rpm) at 15 °C, 12:12 (L:D) and a sinusoidal light regime with a maximum irradiance of 80 μmol photons·m^{-2}·s^{-1} (Fig. 3.3 to 3.6). Once this protocol was established, it was deployed to maintain all the available strains ($N_{STRAINS}$ total = 644 strains) and the survival rate of these strains was recorded over 15.5 months (Tables 3.3 to 3.6).

The survival rate of small *Pseudo-nitzschia multistriata* strains ($N_{STRAINS}$ total = 535 strains) ranged between 49.7 and 97.9% per month over 15.5 months. The lower value may seem disappointing, but the figures take into account all problems encountered. The main factors affecting the survival of strains were of a technical

nature, including agitator overheating, climate chamber malfunctioning and power cuts (Table 3.3). The maintenance procedure has been applied also on initial cells of *P. multistriata* (*i.e.*, large cells at the beginning of the vegetative cycle). Their survival rate monitored over 5 months ranged between 60.0 to 95.9% (including technical problems) (Table 3.4). The procedure was also performed on several other *Pseudo-nitzschia* species (*i.e. P. delicatissima* (clades 1 and 4), *P. pseudodelicatissima, P. mannii, P. dolorosa, P. calliantha,* and *P. galaxiae*) and on *Chaetoceros danicus.* Their survival rate (Table 3.5) ranged from 51.4 to 94.7% over 4 months, species non specific (including technical problems, Table 3.6).

Taking into the account only the small *P. multistriata* strains and the percentage of strains alive (Table 3.3), a theoretical estimation of survival over a large culture batch has been conducted using 1000 strains (Table 3.6). Three conditions have been settled: (i) cultures with maximum efficiency [91.6-97.9 %], (ii) cultures encountering some minor technical problems [67.4-97.9 %], (iii) cultures encountering major technical problems [53.7-89.5 %]. In the first case, up to 61 % of the strains remain alive; the second case *ca.* 30 % and the third case 7.2 %. To reduce the impact of such problems, back-up cultures should be maintained in a separate culture chamber than the stocks and, if possible, strains that present difficulties to grow, be replicated in the maintenance culture.

At each transfer, the survival rate was higher in the stock cultures (new culture well, sinusoidal light) than in the back-up cultures (refreshed cultures, on-off light). Over 6 months, the mortality rate spanned from 19.8 to 45.8% in the stock culture for 18.6 to 92.1% in the back-up culture (Table 3.7).

Table 3.3: Efficiency of the maintenance procedure tested on Pseudo-nitzschia multistriata *over 15.5 months ($N_{STRAINS\ TOTAL}$ = 535). The sequential number of transfers (Tr), the date of each transfer (Transfer date), the number of weeks between transfers (# Weeks), the number of alive strains at each transfer (# Strains, $N_{STRAINS\ MAX}$ = 240 strains), the number of new strains added to the stock (# Entry), the number of dead strains at each transfer (# Dead), the number of cultures that survived in the back-up culture and allow the replacement of dead strains in stock culture (# Recovered), the percentage of alive strains (% Alive) during the procedure and the signalisation of potential problem (P) in the growth chamber (agitator warming or stopped (*), power cuts (Δ), agitator substitution (∞)).*

Tr	Transfer date	# Weeks	# Strains	# Entry	# Dead	# Recovered	% Alive	P
1	17/12/2008	3.01	237	49	52	6	78.1	*
2	13/01/2009	3.09	240	0	5	4	97.9	
3	04/02/2009	3.01	239	12	25	3	89.5	*
4	25/02/2009	3.00	229	0	19	6	91.7	
5	20/03/2009	3.03	214	0	15	4	93.0	
6	14/04/2009	3.06	203	3	13	0	93.6	
7	12/05/2009	4.00	193	7	97	7	49.7	*
8	09/06/2009	4.00	110	0	17	7	84.5	*/Δ/∞
9	02/07/2009	3.03	100	72	8	3	92.0	
10	27/07/2009	3.06	167	0	14	8	91.6	
11	17/08/2009	3.00	161	24	12	10	92.5	
12	16/09/2009	4.03	183	0	26	0	85.8	Δ
13	09/10/2009	3.03	157	139	42	10	73.2	
14	09/11/2009	4.04	264	37	86	1	67.4	
15	03/12/2009	3.04	216	0	100	1	53.7	*
16	07/01/2010	5.00	117	0	40	5	65.8	*
17	01/02/2010	3.06	82	-	16	6	80.5	∞

Table 3.4: Efficiency of the maintenance procedure tested on large cell size cultures (i.e., cultures established from initial cells emerging from the auxospore) of Pseudo-nitzschia multistriata *over 5 months ($N_{STRAINS}$ total = 68). The sequential number of transfers (Tr), the date of each transfer (Transfer date), the number of weeks between two transfers (# Weeks), the number of alive strains at each transfer (# Strains, which is calculated as follows: $N_{STRAINS\ (Tr=n)} = N_{STRAINS\ (Tr=n-1)} - N_{DEAD\ (Tr=n-1)} + N_{RECOVERY\ (Tr=n-1)} + N_{ENTRY\ (Tr=n-1)}$), the number of strains newly added to the stock (# Entry), the number of strains dead at each transfer (# Dead), the number of cultures that survived in the back-up culture and allow the replacement of dead strains in stock culture (# Recovered), the percentage of alive strains (% Alive, which is calculated as follows: % Alive = $(1-N_{DEAD}) \times 100 / N_{STRAINS}$) during the procedure and the signalisation of potential problem (P) in the growth chamber (agitator warming or stopped (*), power cuts (Δ), agitator substitution (∞)).*

Tr	Transfer date	# Weeks	# Strains	# Entry	# Dead	# Recovered	% Alive	P
1	09/10/2009	3.03	40	17	16	7	60.0	
2	09/11/2009	4.04	48	0	12	0	75.0	
3	03/12/2009	3.04	37	0	8	1	78.4	*
4	07/01/2010	5.00	30	22	4	1	86.7	*
5	01/02/2010	3.06	49	-	2	2	95.9	∞

Table 3.5: Efficiency of the maintenance procedure tested on different diatom cultures (i.e. Pseudo-nitzschia *species and* Chaetoceros danicus*) over 4 months ($N_{STRAINS}$ total = 41). Table legend see Table 3.4.*

Tr	Transfer date	# Week	# Strains	# Entry	# Dead	# Recovered	% Alive	P
1	09/10/2009	3.03	19	1	1	1	94.7	
2	09/11/2009	4.04	20	12	9	1	55.0	
3	03/12/2009	3.04	24	16	9	4	62.5	*
4	07/01/2010	5.00	35	-	17	2	51.4	*

Table 3.6: Estimated number and percentage of alive and dead strains over a period 5-7 months cultures considering a massive initial culture ($N_{STRAINS}$ = 1000 strains) under three conditions: (i) maximum maintenance efficiency [91.6-97.9%]; (ii) maintenance efficiency with minor problem encounter [67.4-97.9%]; (iii) maintenance efficiency considering a procedure with only technical problems [53.7-89.5%]. First row indicates the maintenance efficiency evaluated over 1000 initial cultures. Third row indicates the percentage of cultures alive over the culture period for which the conditions (i-iii) were monitored (second row).

Maintenance Efficiency	[91.6-97.9%]	[67.4-97.9%]	[53.7-89.5%]
Number of months in culture	5	7	7
% Alive	61.0	30.1	7.2

Table 3.7: Efficiency of the maintenance procedure on stock and back-up culture conditions. The date of each transfer, the number of dead culture in the stock and in the back-up procedure, the total number of cultures, the percentage of dead cultures in stock and in back-up procedure and the signalisation of potential problem (P) in the growth chamber (agitator warming or stopped (*), power cuts (Δ), agitator substitution (∞)).

Date	Stock	Back-up	Total	% stock	% back-up	P
16/09/09	17	16	86	19.8	18.6	Δ
09/10/09	40	83	117	34.2	70.9	
09/11/09	98	197	214	45.8	92.1	
03/12/09	27	104	134	20.1	77.6	*
07/01/10	40	77	155	25.8	49.7	*
01/02/10	42	93	183	23.0	50.8	∞

3.4. Discussion

The protocol established here (Fig. 3.3 to 3.6) permits reliable long-term maintenance of *P. multistriata* strains in 12-well plates. The main steps include (i) inoculation of fresh medium with a low number of cells (*ca.* 76), (ii) incubation under continuous orbital agitation (70 rpm) at 15 °C, 12:12 (L:D) and a sinusoidal light regime with a maximum irradiance of 80 μmol photons$\cdot m^{-2} \cdot s^{-1}$ and (iii) a weekly maintenance of all strains. Such a procedure regime deployed to 200 strains implies: a transfer of strains every four weeks, *ca.* eight hours of work every four weeks (2:30 hours of preparation (Fig. 3.3), 4:30 hours of culture transfer (Fig. 3.4 and 3.5)) and a weekly brief control (15 minutes; Fig. 3.6). The disposables needed for 200 strains include 1200 ml of F/2 filtered medium, 17 12-well-plates, 200 Gilson pipette tips, 200 plastic mono-use pipettes and Parafilm.

This protocol, or at least the strategy behind establishing this protocol, may help culture collections enrich their holdings of diatom strains, and with those of *Pseudo-nitzschia* species in particular. The protocol has three advantages. First, maintenance effort, as measured in person hours per strain per year, decreases markedly. The use of multi-well plates instead of bottles is an important factor herein because it drastically reduces the number of individual manipulations. Second, the space required to maintain strains decreases because each strain needs only one well in a 12-well culture plate and one well in a back-up plate instead of a pair of bottles. The third and most important advantage is that strains can be kept over extended periods in a reliable way. If a stock culture dies or is not transferred properly then there still is the back-up. It is recommended to keep the back-up plate stored in a different culture chamber to reduce losses resulting from technical

malfunctioning. All these advantages are significant issues for commercial and scientific culture collections.

An alternative way to solve the problem of a restricted lifetime of diatom strains is to let them have sex and maintain the subsequent F_N generations. For many applications, however, this is not a solution, as it requires conservation of strain identity. In addition, inducing sexual reproduction is not easy in many diatom species. It requires maintenance of several strains of each biological species, especially in those exhibiting different mating types, and it requires knowledge on how to induce sexual reproduction in the species at hand, which is unknown for most of the diatoms (Chepurnov et al., 2004).

The exploration of optimal conditions for growth has been conducted in series, by optimising one parameter after another. In the case presented here, first temperature, then irradiance intensity and light regime, then agitation mode and then inoculum density, taking into account pre-existing experience and knowledge (bibliography) as well as clues obtained from field conditions in which *P. multistriata* was encountered as guides of which factors were the most relevant to explore. To determine the relevant parameters to explore, I took into the account the field conditions (EEP-data, (D'Alelio et al., 2010)) during and outside the period in which *Pseudo-nitzschia multistriata* is encountered in the phytoplankton samples taken at the LTER-MC.

The chain of optimization steps implies that when for instance a proper temperature was determined for slow growth, the other parameters had to be chosen based on literature data and guesses. So, the light regime in this experiment was 'on-off' and no shaker was used. Later in the procedure it was discovered that a sinusoidal light regime and gentle orbital shaking greatly improved culture long-term

survival. The disadvantage of such a strategy is that factors may affect the optimal and ranges of other factors, meaning that one may miss the optimum by far and one may neglect or miss the importance of certain factors, thereby never reaching the absolute optimum.

In the present study, temperature and irradiance were the first potentially growth-limiting factors to be explored, as they are known to have an overarching effect on the growth rate. Churro *et al.* (2009) developed a protocol in which they exposed *Pseudo-nitzschia species (P. americana, P. australis, P. calliantha, P. cuspidata, P. delicatissima, P. dolorosa, P. fraudulenta, P. pseudodelicatissima, P. subpacifica* and a new variety of *P. pungens,* and *P. multistriata*) to 15-20 °C but used their protocol for relative short-term maintenance of cultures. The temperature they deployed is close to that encountered in the seawater during the period in which their species were found to bloom (12-19 °C ± 2). These authors may not have explored if somewhat lower temperatures could have permitted longer incubation-periods. D'Alelio *et al.* (2009b) and Amato *et al.* (2010) compared the growth rate, reduction size and domoic acid production (DA) between *P. multistriata* strains from the Gulf of Naples (Italy) growing at two different conditions: (i) low light and low temperature and (ii) high light and high temperature. D'Alelio *et al.* (2009b) showed a higher growth rate (1.27-2.40 versus 0.26-0.29 divisions·day^{-1}) at high temperature and high light (150 μmol photons·m^{-2}·s^{-1} at 25 °C) than at low temperature and low light (10 μmol photons·m^{-2}·s^{-1} at 15 °C). In addition, D'Alelio *et al.* (2009b) uncovered that the reduction rate of the cell length depended on the cell length (*e.g.* the taller the cell, the higher the growth rate and the size reduction). The average reduction of the cell length per division is much higher in the beginning of the lifetime of a strain, following its emergence of the initial cell from the auxospore, than later in its lifetime.

These results indicate that low temperatures (15 °C) are better to decrease the growth rate and to maintain a culture over a long period.

Day-length was not adjusted because this factor does not vary radically over the annual cycle in the Gulf of Naples and 12:12 L:D is the photoperiod assuming neutral for the species. However a comparison of a sinusoidal light regime and an on-off regime was justified because the former reflects natural conditions, and indeed, cultures fared markedly better. The effect of mixing was assessed to reproduce phytoplankton environmental conditions, which present very mild but complex water movement. A complete absence of water movement is by definition unnatural and therefore might be experienced as stressful, and indeed, mild agitation markedly improved survival. The only reason for not recommending more rigorous mixing regimes is that they can lead to cross contamination among wells in plates due to spill over.

An incubation period of *ca.* four weeks between subsequent transfers was targeted because a procedure of monthly transfers with weekly checks is easy to plan into a laboratory routine. The maintenance protocol ensured a period of *ca.* three weeks of slow exponential growth and a slowing down of growth during the remainder of the time without cultures becoming moribund. Amato *et al.* (2010) showed that when strains of *P. multistriata* were incubated in F/2 medium with low P, N or Si concentrations, growth was initially as fast as in normal medium but it slowed down earlier because of an earlier depletion of the critical element in the medium, and hence to an earlier onset of culture collapse, which is, of course, precisely what has to be avoided when cultures are to be maintained as long as possible. Therefore, full strength F/2 medium was used to postpone the onset of nutrient limitation and to thus extend the period between subsequent culture transfer dates. The length of

time, over which exponential growth was observed, depended also on the number of cells inoculated. The lower that number is, the longer it takes the growing culture to deplete its resources. In theory, one could transfer just one or a few cells, but that becomes impractical and may increase the risk of accidentally transferring none.

The procedure has been settled for 200 strains. Results showed that increasing the number of strains per agitator (*e.g.* $N_{STRAINS}$ = 481 the 09/11/2009) induce a decrease in the maintenance efficiency. One solution should have been the use of two agitators per culture chamber (stock and back-up): less shadowing effect of neighbouring piles, less condensation on the plate wall, therefore increasing the light availability and the long term efficiency of culture maintenance. Actually, when the number of strains per agitator was restricted to about 200 strains ($N_{STRAINS}$ max = 240; Table 3.3, Tr 1 to 12) the culture efficiency was higher (91.6-97.9%, without major technical problems).

The results also suggest that genetic factors also affect the growth rate under given environmental conditions. Differences in growth rates imply that the maintenance protocol will select for the slower growing strains, as the faster ones have a higher propensity to perish. For some application this may not be a problem, but for population genetic approaches all sampled strains must be included in the investigations. Moreover, micro-evolutionary events, however low, may occur in a somatically dividing culture (Chapter 8). The transfer of a very low number of cells may increase genetic drift, through founder effects, in the strain culture over prolonged periods, thus altering strain identity over time. The number of cells in inoculates may also explain the higher mortality in back-up cultures (60% over 6 months) because back-up cultures were started with a higher number of cells than for stock cultures.

Antibiotics could be deployed to arrest bacterial overgrowth in *P. multistriata* cultures, but it is questionable. Short-time exposure of cultures of *P. multistriata* to antibiotics revealed that antibiotics reduced the robustness of cells and affected the growth rate of cultures negatively and erratically (personal communication, Dr Lüdeking). The reason for this negative effect might be the existence of functional links between *Pseudo-nitzschia* species and bacteria (Bates *et al.*, 2004; Kaczmarska *et al.*, 2005). The epibacteria of *P. multiseries* belonged to Alpha- and Gammaproteobacteria and Bacterioides (Kaczmarska et al., 2005). However the nature of the interaction between bacteria and *Pseudo-nitzschia* species is not yet solved: (i) phytoplankton may use toxin, as metal chelator, to increase ion up-take (Rue *et al.* 2001) and evolved bacteria in allelopathic communications (Cembella, 2003), (ii) bacteria may increase phytoplankton allelopathic production (Uribe & Espejo, 2003) or (iii) bacteria may decompose waste products and secondary compounds exuded by the diatom cells into the medium (Bell, 1983; Hulot & Huisman, 2004; Obernosterer & Herndl, 1995; Sell & Overback, 1992), therefore the increase of bacteria and mucilage may harm the diatoms when concentrations become too high (Obernosterer & Herndl, 1995); (iv) bacteria may produce secondary compounds that are essential for growth and survival of the diatoms (*e.g.* endo-symbiotic bacteria in axenic-diatom culture). For example, bacteria provide vitamins essential for normal growth in macroalgae (Croft *et al.*, 2006). At least in macroalgae, growth ceases or becomes callus-like under axenic conditions, but re-establishes itself immediately upon re-introduction of bacteria associated to not axenic cultures (Kooistra *et al.*, 1991); (v) bacteria may compete for essential nutrients in the growth medium.

Antibiotics were not used in the current study, as axenic cultures require elaborate procedures of transfer and maintenance. Mild agitation suppressed the proliferation of bacteria but the reason why bacteria proliferate especially in the absence of mixing has not been investigated. It might be that, in the absence of mixing, diatoms sink to the bottom of the culture vessel where they form unnaturally high concentrations, depleting resource and releasing exudates, or establishing other conditions that favour the growth of bacteria.

The protocol presented here, has been applied to cultures of *P. multistriata* strains early in their lifetime, *i.e.*, immediately upon strain establishment from large initial cells. It has been applied also to several other *Pseudo-nitzschia* species and to *Chaetoceros danicus*. All the tested strains were maintained for extensive periods, exceeding by far their lifetime in the lab in comparison to maintenance regimes designed to foster rapid growth. The procedure developed here, however, is unlikely to be universally applicable for all diatoms. For example, diatoms collected from tropical areas might require higher minimum temperatures to ensure slow growth and reliable survival over extended periods, whereas those from cool-temperate and polar regions might tolerate temperatures close to freezing of seawater for extremely slow but extended growth.

Culture maintenance summary in three phases:

(i) Preliminary steps (Fig. 3.3), (ii) Transfer steps (Fig. 3.4 and 3.5) and (iii) Weekly maintenance (Fig. 3.5)

***Preliminary steps*: (ca. 2:30 hours)**

1) Control of "stock" well under light microscope (100x magnification).

2) Manage list of strains present in the stock cultures and identify the transfer strategy (retrieve perished stock strains from back-up, if alive and well there).

3) Filtration of F/2 medium.

4) Preparation of the new stock-plates (rinse them, write information on lid).

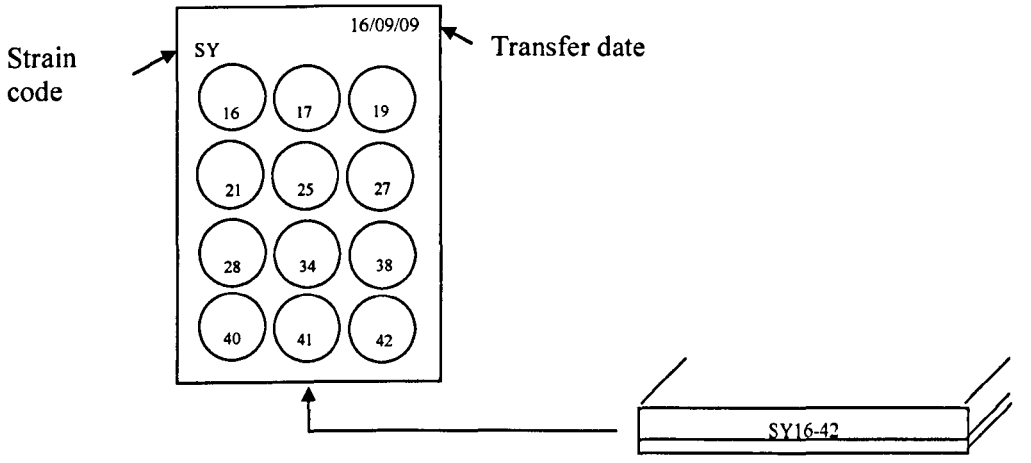


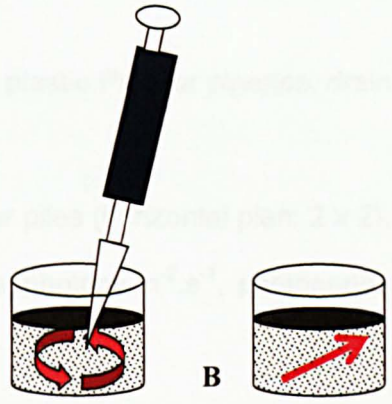
Figure 3.3: Outline of the preliminary steps in the culture maintenance procedure. The cartoon illustrates the fourth step: preparation of the new stock-plates.

Figure 3.4: Outline of the Transfer steps in the maintenance procedure illustrating the establishment of new stock culture. The cartoons represent the homogenisation step (A), the sampling transect (B), the inoculation transfer (C) and the placement of piles of plates on the orbital shaker (view from the top) (D).
(please see next page)

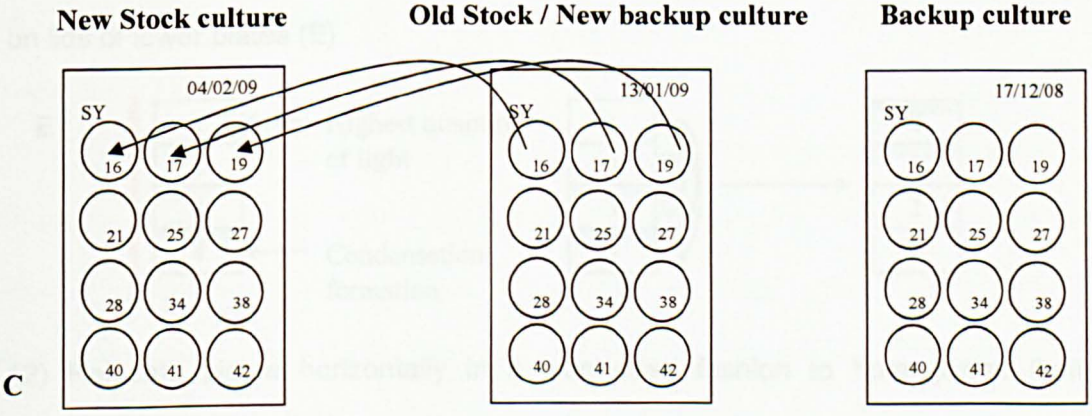
***Transfer procedure*: (*ca.* 4:30 hours)**

- Stock cultures:

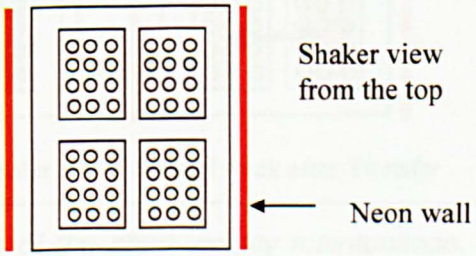
5) Homogenise stock culture with Gilson pipette tip (Uniform and centric homogenisation (**A**) to detach cells from the well bottom and homogenise the culture)



6) Transfer 1.2 µl (categories 1-3) to 2.4 µl (categories: 4-5) of homogenised culture medium from the old stock-plate to the new stock-plate (**C**) (Inoculate taken in a vertical-horizontal transect (**B**) from bottom to top and part to part of the culture well).



7) Seal new stock culture plate with Parafilm and place in four piles (2 x 2), at 15°C under sinusoidal light (light impulse at 40 *µmol* photons.m^{-2}.s^{-1}, sinusoidal curve with a maximum of 80 *µmol* photons.m^{-2}.s^{-1}), photoperiod 12:12 L:D, on orbital agitator (70 rpm).



- Back-up cultures:

8) Once transferred, homogenise stock cultures with plastic Pasteur pipettes, drain and refresh with 2 ml of F/2 medium.

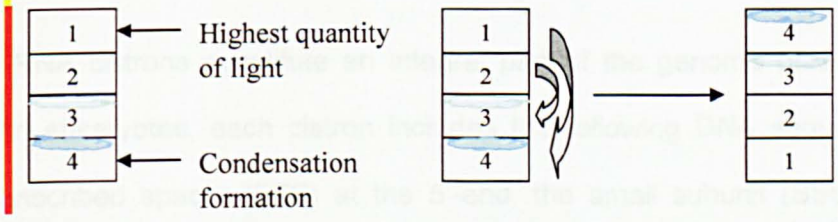
9) Seal back-up plates with Parafilm and place in four piles (horizontal plan: 2 x 2), at 15°C under sinusoidal at a maximum of 80 *µmol* photons.$m^{-2}.s^{-1}$, photoperiod 12:12 L:D, on the orbital agitator (70 rpm).

10) Discard old back-up plates.

Figure 3.5: Outline of the steps to establish the new back-up cultures.

***Weekly maintenance*: (ca.15 minutes)**

11) Move plates vertically from bottom to top to counter condensation deposition on lids of lower plates (**E**).



12) Relocate plates horizontally in a clock-wise fashion to homogenise light capture from vertical neon panel, after circa four weeks of incubation (**F**).

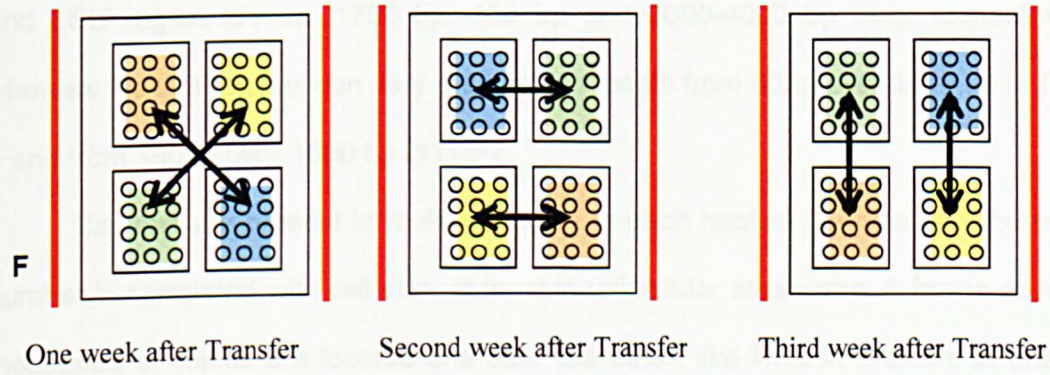


Figure 3.6: Outline of the short weekly maintenance. The cartoons represent the vertical (E) and horizontal (F) migration of culture plates, once a week.

CHAPTER FOUR

DIVERSITY OF THE INTERNAL TRANSCRIBED SPACER REGIONS ACROSS BLOOMS OF *PSEUDO-NITZSCHIA MULTISTRIATA* IN THE GULF OF NAPLES

4.1. Introduction

In the present study intraspecific and intra-individual variation in the internal transcribed spacers (ITS-1 and ITS-2) within the nuclear ribosomal RNA (rRNA) cistrons are used to assess population genetic structure in the planktonic pennate diatom species *Pseudo-nitzschia multistriata* in the Gulf of Naples. Since the use of the ITS region for this particular purpose is unusual, I will explain below why and how I used it.

The RNA cistrons constitute an integral part of the genome of every living organism. In eukaryotes, each cistron includes the following DNA sequences: an external transcribed spacer (ETS) at the 5'-end, the small subunit (SSU or 18S) rRNA-coding gene, the ITS-1, the 5.8S rRNA-coding gene, the ITS-2, the large subunit (LSU or 28S) rRNA-coding gene and an ETS at the 3'-end. The SSU, 5.8S and LSU region are *ca.* 1750 bp, 180 bp and 3000-4000 bp long, respectively, whereas the ITS regions can vary markedly in length from 50 to over 1000 bp in ITS-1 and from 180 to over 1000 bp in ITS-2.

Cistrons are present in multiple copies in each haploid genome and the copy number is correlated with cell size, at least in unicellular organisms. A few to several thousands of copies are located one after the other, like links in a chain, in one or more locations in each haploid set of chromosomes. In Bacillariophyceae, estimates

of copy numbers range between about 40 and 600 (Creach *et al.* 2006, Zhu *et al.* 2005, Fig. 4.1).

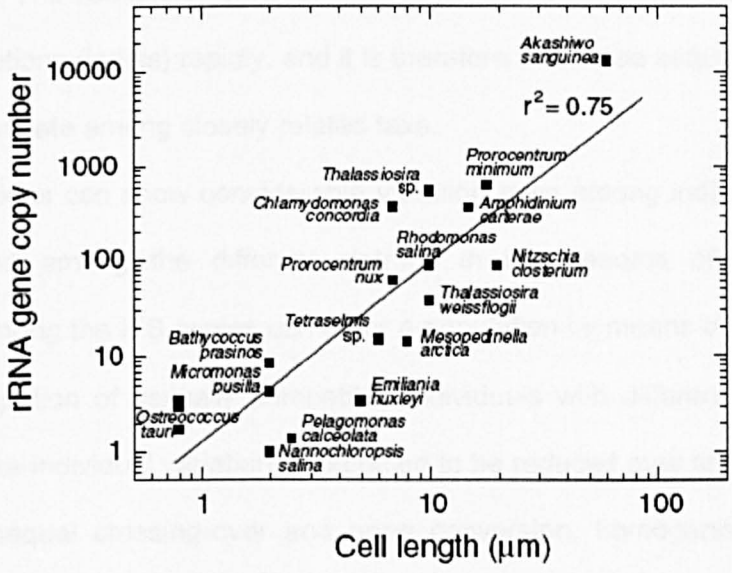


Figure 4.1: Estimation of rRNA gene copy number in function of cell length over 18 phytoplanktonic strains (Zhu et al. 2005).

The whole cistron is transcribed into RNA. As the ribosomal RNA (rRNA) molecule is formed, it folds into a characteristic secondary structure by means of complementary sequence regions forming double-stranded stems, with single stranded loops in between. After the spacers splice themselves out of the RNA molecule, the remaining three RNA components (SSU rRNA, 5.8S rRNA, LSU rRNA) group together with proteins to form a mature ribosome. The ribosomes constitute the protein factories of the cell, translating messenger RNAs into proteins (Lafontaine & Tollervey, 2001).

The SSU rRNA, 5.8S rRNA, and LSU rRNA are functionally conserved in both length and way of folding, and therefore accumulate substitutions relatively slowly, although site-specific differences in substitution rates occur. This is why these regions are used to infer relationships among genera and higher taxa. Instead, the

ITS regions are less functionally conserved because they are not incorporated in the ribosome. They just need to fold properly, which is performed by relatively conserved core sections. The remaining parts of the ITS, accumulate base substitutions and insertions-deletions (indels) rapidly, and it is therefore that these sequences are often used to discriminate among closely related taxa.

ITS regions can show considerable variation even among individuals within a population and among the different cistrons in the genome of an individual. Differences among the ITS copies can enter a population by means of mutations and through immigration of sexually compatible individuals with different ITS-types. All this arising intra-individual variation is expected to be reduced over time because two processes, unequal crossing-over and gene conversion, homogenise the various cistron copies, fixing one or another ITS-type or a recombinant resulting from crossing-over inside a cistron or even inside the ITS-region.

Intraspecific ITS variation can render reading of electropherograms generated by sequencing PCR-products of these ITS regions cumbersome. Base differences (for example A - G) at single sites (single nucleotide polymorphisms, SNPs) in different ITS copies appear as double peaks in the electropherogram of a sequence read, and are denoted as ambiguities (R = A+G). Other ambiguities are denoted as Y (= C+T), M (= A+C), K (= G+T), S (= C+G), W (= A+T), B (= non A), D (= non C), H (= non G) and V (= non T, or U in the case of RNA) (Cornish-Bowden, 1985). Ambiguities because of indels are more difficult to read because from the position at which the indel starts towards the 3-'end the electropherogram is composed of two reads of the same sequence, but offset by the length difference of the indel. The electropherogram shows ambiguities where the two bases in the offset sequence

positions happen to be different, and it reveals single peaks where the bases happen to be the same. I refer to such patterns from here on as double reads.

A way to unravel this intra-individual variation is to clone the PCR product of the ITS region and sequence, say, 40-50 DNA-clones. The finding of only two distinct ITS-types in the clone library suggests that the individual is a hybrid between two genetically distinct parents. If recombinants between the ITS types are found as well, then the individual is probably the progeny of hybrid parents and the recombinant ITS sequences have been generated by crossing-over between the two ITS-types in hybrid parents, though recombinants could be PCR artefacts as well. However, cloning is financially prohibitive when hundreds of strains are examined.

An alternative to cloning is to interpret the double reads in electropherograms, but that is feasible only when the length of the indel and the sequence beyond the indel are known. D'Alelio *et al.* (2009a) observed two ITS sequence types in strains of *P. multistriata* sampled from the Gulf of Naples (GoN), which these authors defined as ITS A-type and ITS B-type. Strains showed either the A-type, or the B-type, or both types (A/B-type; Table 4.1). ITS A-type and ITS B-type differed as follows from the 5'-end to the 3'-end (Table 4.1):

<u>On ITS-1</u>:

- A SNP: a T in the ITS A-type and a C in the ITS B-type;
- Immediately thereafter, a microsatellite CATTG: absent in the ITS A-type and present in the ITS B-type;
- Immediately thereafter, a microsatellite TATTG: absent in the ITS A-type and present one to more fold in the ITS B-type;

- Immediately thereafter, a microsatellite TA: twice in the ITS A-type and twice or more in the ITS B-type;
- Immediately thereafter, a SNP: a C in the ITS A-type and a T in the ITS B-type;
- A SNP: a T in the ITS A-type and a C in the ITS B-type.

On ITS-2:

- An insertion ACTTTACTACTA: absent in the ITS A-type and present in the ITS B-type,
- A microsatellite ACT: twice in the ITS A-type and once in the ITS B-type.

	ITS-1							ITS-2	
Aligned Position	69	131	141	146	151	170	301	579	793
Variation	SNP	(TATTG)x	CATTG	(TATTG)x	(TA)x	SNP	SNP	Indel	(ACT)x
A	T	2	-	-	2	C	T	-	2
B	C	2	1	1 or more	2 or more	T	C	1	1
A/B	T>C	2	1	1 or more	MIX	MIX	MIX	MIX	MIX

Table 4.1: Representation of ITS-1 and ITS-2 rDNA region differences from direct sequencing as a function of the ITS-(A, B, or A/B) type to which the strain belongs (reading frame sense: 5' to 3'). First row indicates the ITS-1 and ITS-2 region. Second row indicates the position of the variation in ITS-1 (green numbers) and ITS-2 (orange numbers). Third row, the type of variation, with "SNP" for single nucleotide polymorphism, "x" for core repetition and "Insertion" (ACTTTACTACTA) for the insert sequence in ITS-2. Rows four to six indicate the DNA sequence for ITS-A, -B and -A/B.

Electropherograms from ITS A-type strains did not show ambiguities. Instead, those from ITS B-type strains possessed two microsatellite regions of variable length, resulting in a double reading pattern in the forward and reverse reads from the microsatellite onwards. However, this did not pose a problem because the reads in

forwards and reverse direction reached the microsatellite (Fig. 4.2). Interpretation of electropherograms from strains of the ITS A/B-type was more challenging. In the forwards reads, double reading patterns commenced in ITS-1 at the single nucleotide polymorphism (aligned position: 69 for T/C substitution, Table 4.1) and the microsatellite immediately thereafter (aligned position: 141 for absent/presence of CATTG, Table 4.1), and in the reverse reads double reading patterns started in ITS-2 from the microsatellite ACT (two in the ITS A-type and one in the B-type). Nevertheless, it is possible to trace both types in the electropherograms using the sequence reads of the ITS A-type and ITS B-type as guides (Fig. 4.2).

D'Alelio et *al.* (2009a) cloned the rDNA ITS region of strains exhibiting the ITS A/B-type and demonstrated the presence of recombinant ITS sequences. They interpreted this as a sign for sexual compatibility between strains of the ITS-A and ITS-B type in the field. Moreover, the presence of recombinant sequences in strains of the ITS A/B-type suggested that the hybrid offspring of crosses between parents of the ITS-A and ITS-B type are fertile. The authors confirmed this interpretation by demonstrating sexual compatibility between ITS-A and ITS-B strains in the lab. Thus, strains of *P. multistriata* with markedly different ITS-types belong to a single biological species (D'Alelio et al., 2009a).

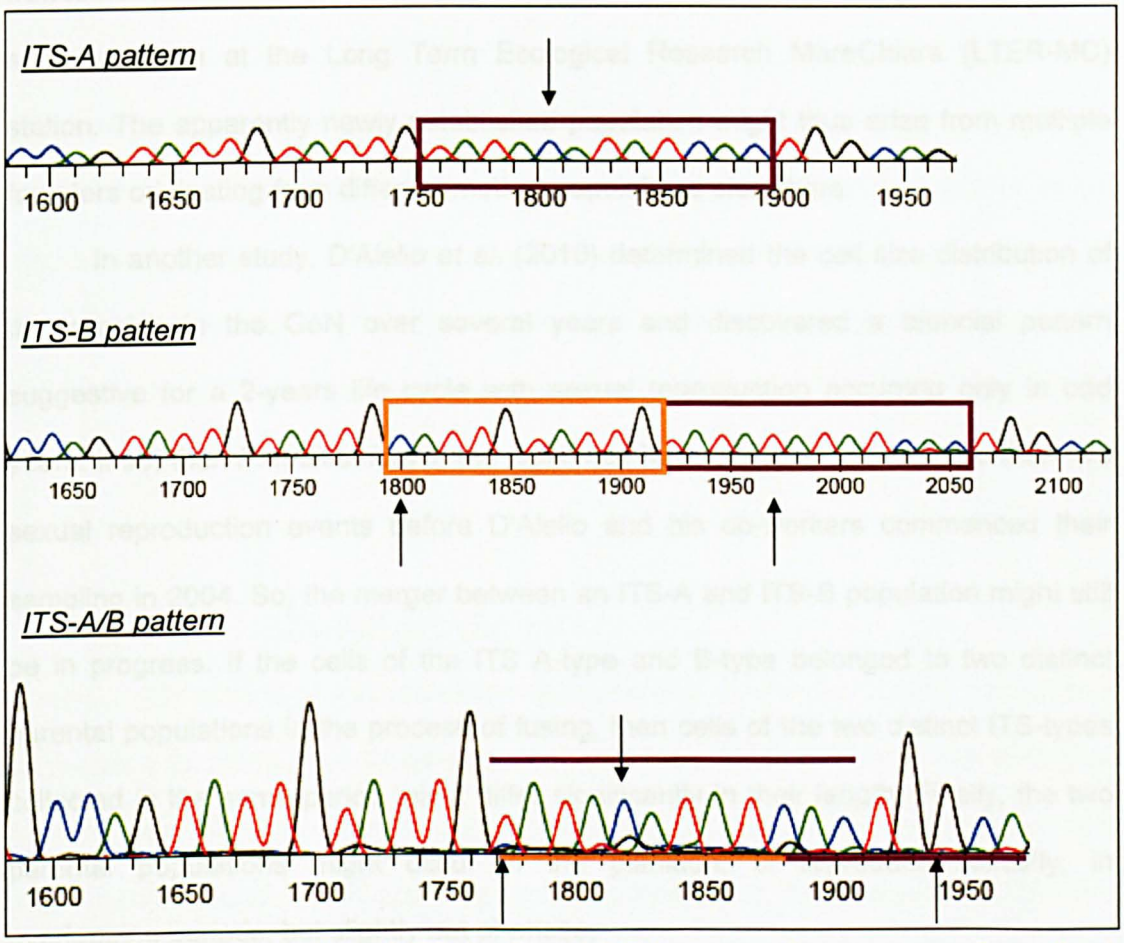


Figure 4.2: Representation of ITS-A, -B and A/B-type genetic sequences from direct sequencing with A (green), T (red), C (blue) and G (black) (Sequence Scanner, version 1.0). Purple box indicates position 141 to 178 of ITS-A type ((TA)x (C/T) ATATCAC). Orange box indicates position 141 to 150 of ITS-B type (CATTG (TATTG)x). Same positions on ITS-A/B type are indicated by purple/orange lines. Arrows indicate the specific base changes.

D'Alelio *et al.* (2009a) suggested that the increasing proportion of strains showing an ITS A/B-type at the cost of diminishing proportions of strains of the pure ITS A- or B-type over the three years of their study result from a merger of two sub-populations of *P. multistriata*, one of the ITS A-type and the other one of the ITS B-type into a single one of a mixed ITS-type. If so then samples in 2008 and 2009 should consist almost entirely of ITS A/B-type cells. A reason why such sub-populations could co-exist in the same place is that the species is a relative

newcomer in the Gulf of Naples; It appeared only as recent as 1995 in the plankton samples taken at the Long Term Ecological Research MareChiara (LTER-MC) station. The apparently newly established population might thus arise from multiple founders originating from different mother-populations elsewhere.

In another study, D'Alelio *et al.* (2010) determined the cell size distribution of the species in the GoN over several years and discovered a biennial pattern suggestive for a 2-years life cycle with sexual reproduction occurring only in odd years. If so, then *P. multistriata* in the GoN may have experienced no more than five sexual reproduction events before D'Alelio and his co-workers commenced their sampling in 2004. So, the merger between an ITS-A and ITS-B population might still be in progress. If the cells of the ITS A-type and B-type belonged to two distinct parental populations in the process of fusing, then cells of the two distinct ITS-types collected in the same period might differ significantly in their length. Finally, the two parental populations might occur in the plankton, or reproduce sexually, in overlapping periods, but slightly out of phase.

To test these hypotheses, 467 strains of *P. multistriata* were sampled in the Gulf of Naples over the early summer and autumn occurrences of the species in 2008 and 2009, measured cell lengths of the strains, and gathered their ITS-types. Sampling of strains encompassed the entire period of appearance of the species to assess if changes in the proportions of strains exhibiting the various ITS patterns occurred over this period. The observed size distribution of cells collected in 2008 and 2009 was compared with predictions by the life cycle model of *P. multistriata* by D'Alelio *et al.* (2010). With sexual reproduction expected to have happened in 2005 and 2007, the 2008 samples should include tall cells of the 2007-cohort and small cells of a parental 2005-cohort. To assess if this 2005-cohort could have given rise to

the 2007-cohort, the ITS-types of the tall and small cells were analysed in 2008 separately.

Cloning the PCR products of 467 strains to obtain the ITS-types would be prohibitively expensive and labour intensive. Therefore the direct interpretation method was used described above to evaluate the proportions of strains of ITS A-type, ITS B-type and ITS A/B-type. Moreover it has been determined for each strain belonging to ITS A/B-type, which of the two ITS types clearly dominated in the electropherogram (ITS A-dominant type, ITS B-dominant type and ITS A/B-mix type).

4.2. Materials and Methods

A total of 467 strains of *Pseudo-nitzschia multistriata* (Appendix 1) were established from cells and chains (see Chapter 2) sampled at the LTER-MC in the Gulf of Naples from July and September to December 2008, from the end of June to the end of July 2009 (early-summer 2009) and September to November 2009 (autumn 2009). For each of the strains the cell length was measured over the apical plane as soon as possible following culture establishment. The ITS region was sequenced successfully for 265 strains: 167 strains from the summer and autumn sampling of 2008 (of which 3 in summer and 164 in autumn), 54 strains from the early-summer sampling of 2009 and 44 strains from the autumn sampling of 2009. Only the results of successful ITS amplifications were therefore presented in this Chapter.

Electropherograms of six strains from the autumn of 2004, of 16 strains from the autumn of 2005 and of 11 strains from the autumn of 2006, generated by D'Alelio *et al.* (2009a), were re-examined to allow identical interpretation procedures for the electropherograms of all strains. Thus, the ITS dataset was established from 298 of the 500 strains gathered in the Gulf of Naples between 2004 and 2009. Strains for

which the ITS pattern was not obtained (PCR-amplification failure or unreadable and / or short electropherogram) have been referred to as ITS-NA (not amplified).

Amplification and sequencing of the ITS regions (ITS-1 5.8S rDNA ITS-2) were carried out as described in Chapter 2. Assignment of the ITS-type was done as follows: a strain was considered to have an ITS A-type if the forward and reverse reads were entirely conform to the description of the ITS A-type of D'Alelio *et al.* (2009a) (see Table 4.1 and Fig. 4.2), and the same accounted for the ITS B-type. Strains were classified as belonging to the ITS A/B-type if the electropherogram of the forward read showed an Y (= C+T) at the first SNP and a double reading pattern at the microsatellite region onwards (*i.e.* from aligned position 141 to onwards, see Table 4.1 and Fig. 4.2) and the electropherogram of the reverse read revealed a double reading pattern from the microsatellite at the 5'-end of the ITS-2 region onwards (Table 4.1).

All strains showing an ITS A/B-type have been distributed over three sub-categories depending on the relative signal intensities of the ITS-A and ITS-B in the electropherograms. Signal strength was assumed to be correlated with the proportions of ITS copies of the A-type and the B-type in the genome. The three sub-categories of ITS A/B-type were A-dominant, B-dominant and AB-mix. Aberrant double reading patterns were detected in electropherograms of a few strains, but these were due to the co-occurrence of contaminants in the cultures, that is, of eukaryote microbes other than *P. multistriata*. Such reads were excluded from the analysis.

Statistical approaches were inferred using the "comparison of proportions" test (called Z-test) over time and strain number. The Z-test follows the Reduced Normal distribution. The observed distributions of ITS-type frequencies (Zc) were compared with the expected ones ($U_{\alpha/2}$), obtained from the Reduced Normal Table. The common proportion (p_c) of ITS-types between a sample 1 and a sample 2 was calculated as the sum of a particular genotype (n = N x p) over the sum of strains analysed (N):

➢ $$p_c = \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}$$

The observed frequency was calculated as follows:

➢ $$Zc = \left| \frac{(p_1 - p_2)}{\sqrt{\frac{(p_c (1 - p_c))}{p_1} + \frac{(p_c (1 - p_c))}{p_2}}} \right|$$

The null hypothesis (Ho) states that there is no difference between the proportions of the ITS-type considered over two periods tested (or between two ITS-types over the same period of occurrence).

- *The parameter set-up:* under bilateral distribution and an error of 5 % (α = 5 %, $U_{\alpha/2}$ = 1.96).

- *The validation of the test:* number of strains tested (n > 30) and its relative proportion (np > 5). If the Zc value is higher than $U_{\alpha/2}$, Ho cannot statistically be accepted under alpha and H1 is validated.

4.3. Results

Table 4.2: ITS-types (A, B, A/B) successfully sequenced and their respective proportion in the different period sampled (from 2004 to 2009; early summer (☼), autumn (✂) and sum of the two sampling periods (2008-2009).

ITS	2004 ✂	2005 ✂	2006 ✂	2008	2009	2009 ☼	2009 ✂
A	3 ; 0.50	4 ; 0.25	1 ; 0.09	2 ; 0.00	14 ; 0.14	9 ; 0.17	5 ; 0.11
B	2 ; 0.33	3 ; 0.19	1 ; 0.09	20 ; 0.13	30 ; 0.31	15 ; 0.28	15 ; 0.34
A/B	1 ; 0.17	9 ; 0.56	9 ; 0.82	145 ; 0.87	54 ; 0.55	30 ; 0.55	24 ; 0.55
Sum of strains	6	16	11	167	98	54	44
A/B (A-dominant)	1 ; 1.00	5 ; 0.56	8 ; 0.89	12 ; 0.08	36 ; 0.67	20 ; 0.67	16 ; 0.67
A/B (B-dominant)	0 ; 0.00	1 ; 0.11	1 ; 0.11	126 ; 0.87	6 ; 0.11	4 ; 0.13	2 ; 0.08
A/B (AB-mix)	0 ; 0.00	3 ; 0.33	0 ; 0.00	7 ; 0.05	12 ; 0.22	6 ; 0.20	6 ; 0.25

Table 4.2 shows the distribution of ITS-types A, B, and A/B and the ITS-A/B subclasses (A-dominant, B-dominant and AB-mix) over the strains sampled in each of six sample periods (2004-2009). The summer and autumn samples of 2008 were pooled in most of the analyses because the summer sample of that year contained only three strains. As sampling effort differed over years (Table 4.2, Fig. 4.3), the proportions of ITS-types were analysed (Table 4.2, Fig. 4.4).

The proportions of ITS A/B-types (including A-dominant, B-dominant and AB-mix) increased in the samples from 2004 to 2006 (Table 4.2, Fig. 4.4) whereas ITS-A and ITS-B decreased markedly. However, these proportions could not be compared statistically because of the low numbers of strains in the categories (Appendices 4-8). From 2008 to 2009 the proportions of ITS-A and ITS-B increased significantly at the cost of the ITS A/B-types (Fig. 4.4). No significant differences were observed between the proportions of ITS-types in the summer and autumn samples. Statistical tests among proportions of the ITS-A were permitted between samples of early

summer and autumn of 2009, and tests between proportions of ITS-B and ITS-A/B among samples of 2008, 2009 early-summer and 2009 autumn (Appendix 4). Other comparisons violated the validity test (Appendix 6) because there were not enough counts in one or more of the ITS-types.

Results revealed a radical shift within the proportions of the three categories of ITS A/B-types between 2008 and 2009 (Fig. 4.4, Appendix 5). The proportion of strains of the ITS-A/B B-dominant diminished drastically between 2008 and 2009 whereas the proportion of those of ITS-A/B A-dominant increased markedly between 2008 and 2009 (Table 4.2, Fig. 4.4). Results of the re-examined electropherograms of the strains screened by D'Alelio *et al.* (2009a) are shown in Table 4.2 and Figs. 4.3 and 4.4. Numbers of these D'Alelio *et al.* strains were too low to permit statistical testing (Fig. 4.3, Appendices 6-7). Nevertheless, the results for 2005 and 2006 showed the following: a complex ITS A/B-types repartition was shown over the nine strains sampled in 2005 (five ITS-A/B A-dominant, one B-dominant and three AB-mix) whereas 2006 showed a large proportion of ITS-A/B A-dominant over the same number of strains (eight ITS-A/B A-dominant and only one B-dominant). Instead, in 2008, of the 145 strains with an ITS A/B-type, 126 were of the B-dominant type (Table 4.2, Fig. 4.4, Appendix 8).

In Table 4.3 the minimum, maximum and average cell length have been given for each of the sample periods and ITS categories. The results show no significant differences in the mean lengths of the cells among the four main ITS-types in any of the collection periods. Yet cells sampled in the early summer of 2009 were, on average, 4 to 5 μm taller than those collected in the autumn of that year. The high number of ITS-NA strains in 2009 was due to technical problems and to PCR-amplification failure of the ITS-region for a large number of strains.

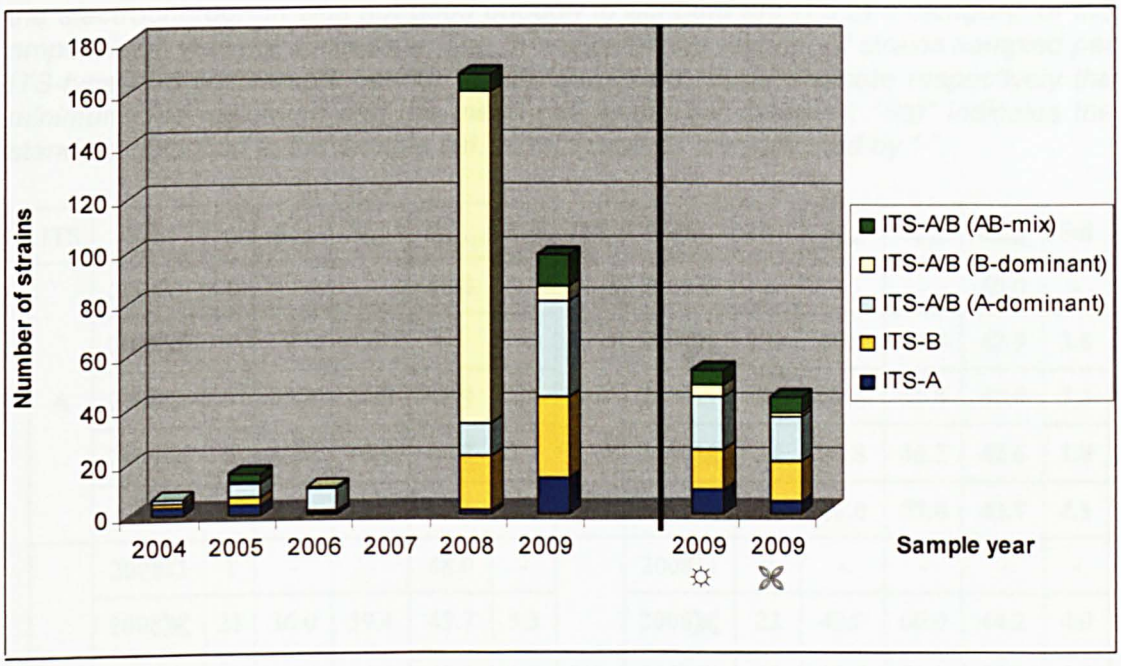


Figure 4.3: Number of strains sampled during the different bloom period (2004-2009; early-summer (☼), autumn (🍂) and sum of the two sampling periods (2008-2009) as a function of their ITS-type.

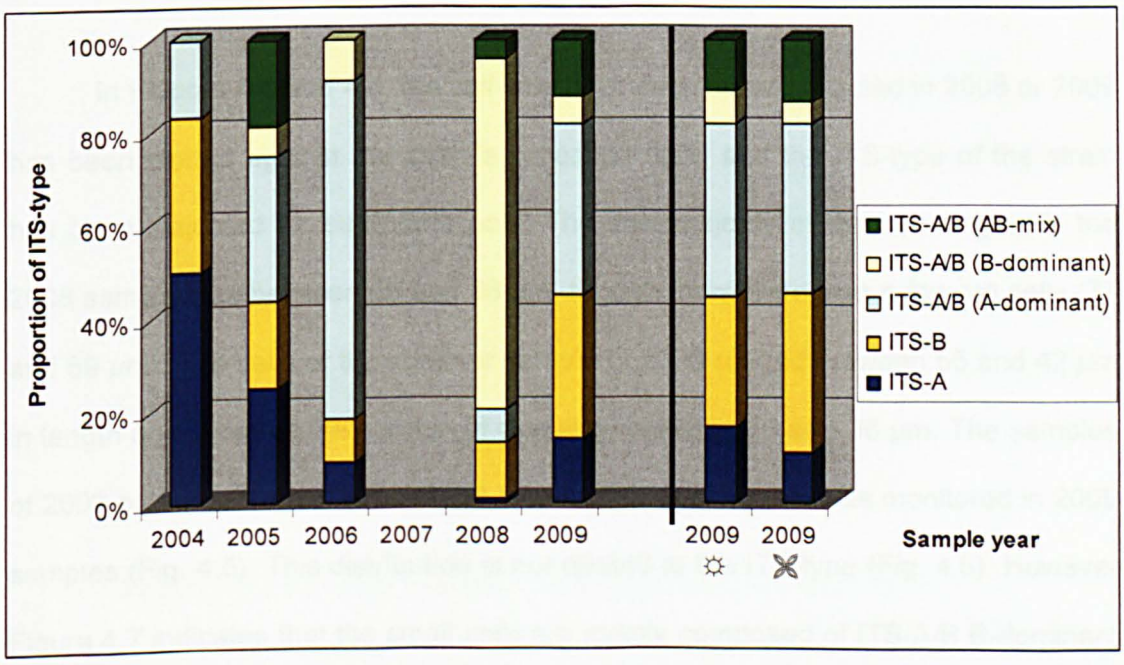


Figure 4.4: Proportion of ITS-types during the different bloom period (2004-2009; early-summer (☼), autumn (🍂) and sum of the two sampling periods (2008-2009).

Table 4.3: Size range per ITS-types (A, B, A/B and Not-Analysed) over early-summer (☼) and autumnal (✂) of 2008 and 2009. NA denotes "ITS type not assignable," i.e., the electropherogram was not good enough to attribute the ITS to a category, or the amplification was not successful. The "n" indicates the number of strains sampled per ITS-type and per sample period; "S_{min}", "S_{max}" and "S_{mean}" indicate respectively the minimum, the maximum and the mean cell length per category; "Std" indicates the standard deviation in the sample set. Missing values are indicated by "-".

ITS	Year	n	S_{min}	S_{max}	S_{mean}	Std	ITS	Year	n	S_{min}	S_{max}	S_{mean}	Std
A	2008☼	1	-	-	48.0	-	A/B	2008☼	1	-	-	60.0	-
	2008✂	1	-	-	-	-		2008✂	138	36.0	77.0	42.9	3.8
	2009☼	9	42.6	54.0	48.9	3.1		2009☼	27	39.6	55.8	47.9	3.3
	2009✂	5	42.0	48.6	44.8	2.5		2009✂	24	37.8	46.2	42.6	1.9
	All	**15**	**42.0**	**54.0**	**47.4**	**3.4**		All	**190**	**36.0**	**77.0**	**43.7**	**4.1**
B	2008☼	1	-	-	48.0	-	NA	2008☼	0	-	-	-	-
	2008✂	21	36.0	59.4	43.7	5.5		2008✂	22	42.0	60.0	44.2	4.0
	2009☼	15	45.0	54.0	48.2	2.8		2009☼	42	49.8	56.4	48.4	3.3
	2009✂	15	41.4	54.0	44.7	3.3		2009✂	135	40.2	54.0	43.7	2.5
	All	**52**	**36.0**	**59.4**	**45.3**	**4.6**		All	**199**	**40.2**	**60.0**	**44.7**	**3.4**

In Figures 4.5 and 4.6, the cell length for each strain collected in 2008 or 2009 has been plotted against the strain's collection date, and the ITS-type of the strain has been indicated for each data point The vast majority of the cell lengths in the 2008 sample was between 50 and 36 µm, though there were also a few tall cells (77 and 59 µm). The cells of the summer sample of 2009 ranged between 55 and 42 µm in length and those of the autumn of that year between 50 and 38 µm. The samples of 2008 presented two groups of cell size whereas one group was monitored in 2009 samples (Fig. 4.5). This distribution is not related to the ITS-type (Fig. 4.6). However Figure 4.7 indicates that the small cells are mainly composed of ITS-A/B B-dominant type; the smallest composed of ITS-A/B A-dominant and ITS-A/B B-dominant cells.

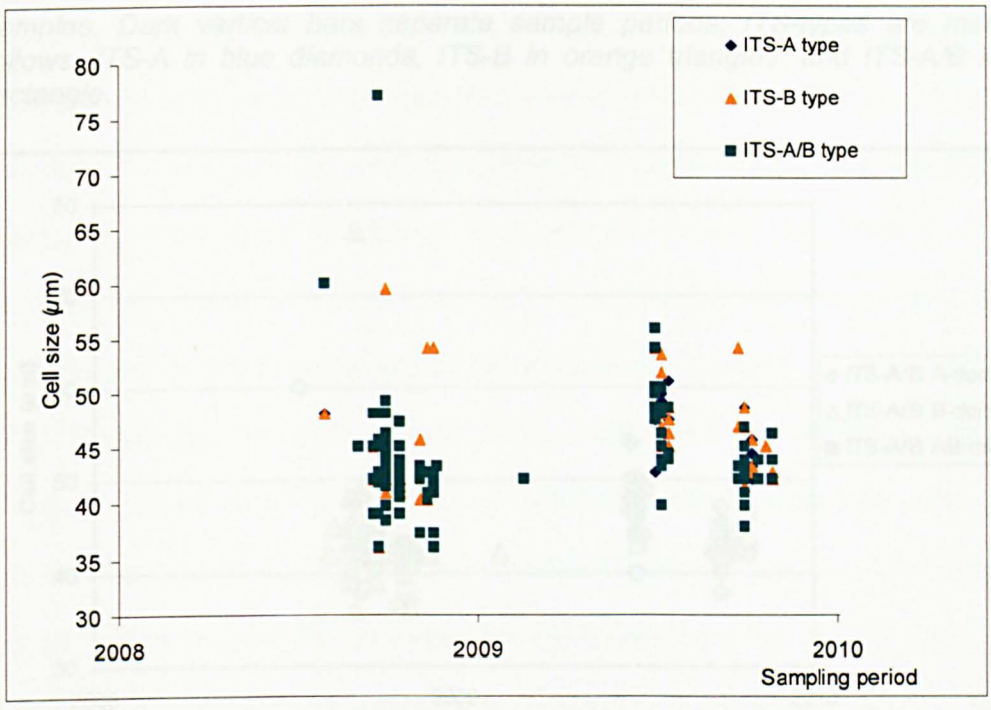


Figure 4.5: Representation of the cell size of 257 strains as a function of their ITS-types, cell size and sampling period (2008-2009; early summer and autumnal blooms). ITS-A is indicated by blue diamonds, ITS-B by orange triangles, and ITS-A/B by green rectangles.

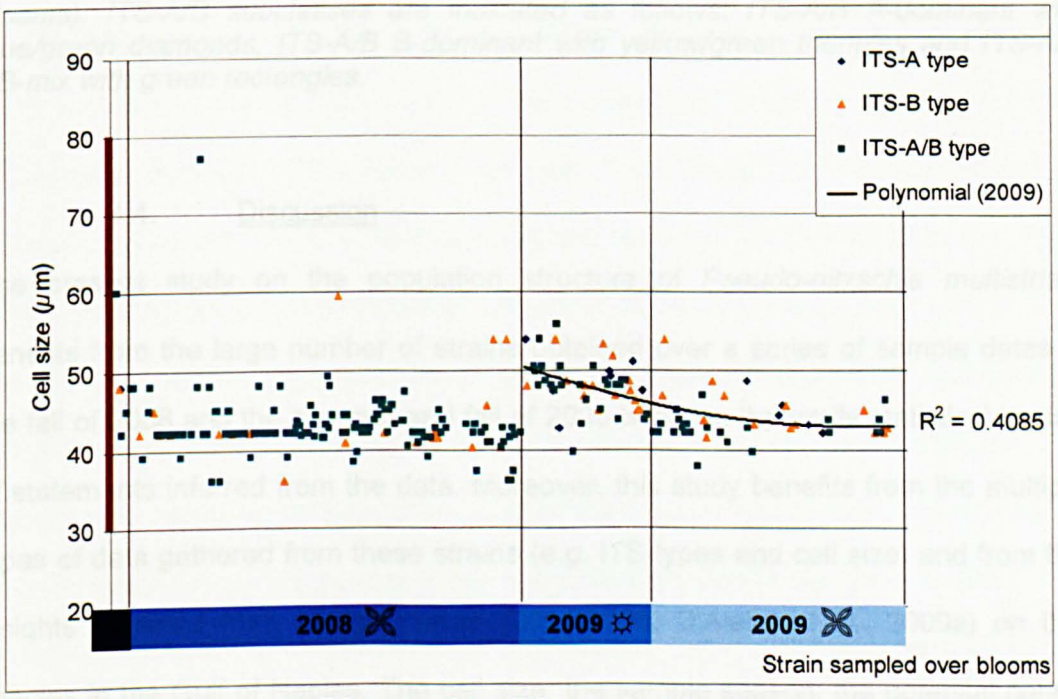


Figure 4.6: Details of the ITS-type of the 257 strains sampling over 2008-2009 bloom periods (early summer (☼) and autumnal (✂) blooms) in function of their cell size. Blue bar graduation indicates 2008☼, 2008✂, 2009☼ and 2009✂ bloom periods. The brown bar indicates the cell size range of P. multistriata observed in field

samples. Dark vertical bars separate sample periods. ITS-types are marked as follows: ITS-A in blue diamonds, ITS-B in orange triangles, and ITS-A/B in green rectangle.

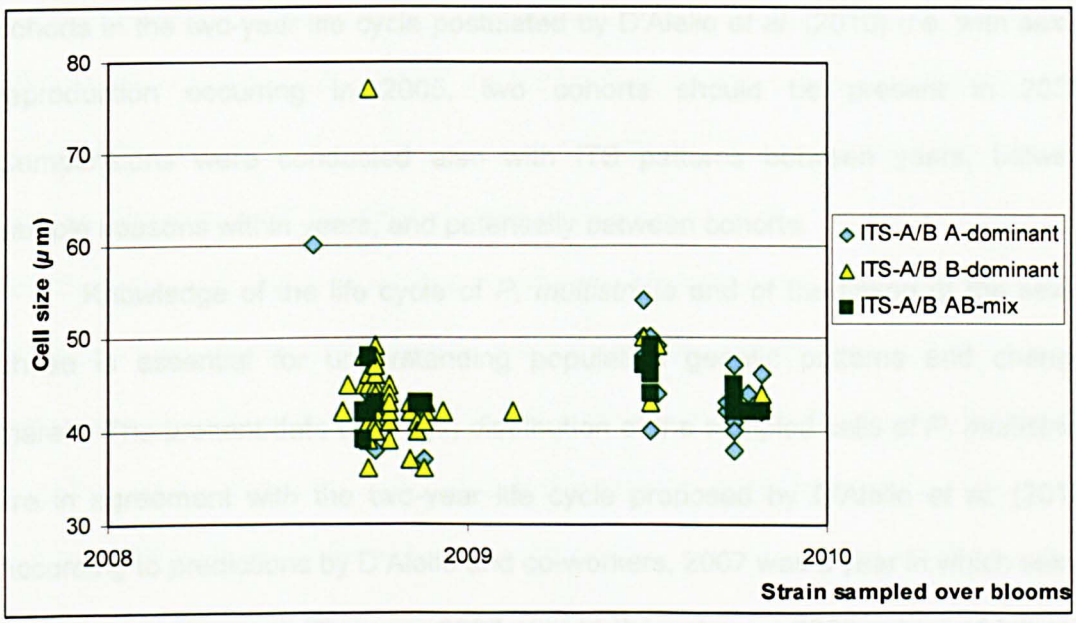


Figure 4.7: Representation of the cell size ITS-A/B strains in function of their ITS-types, cell size and sampling period (2008-2009; early summer and autumnal blooms). ITS-A/B subclasses are indicated as follows: ITS-A/B A-dominant with blue/green diamonds, ITS-A/B B-dominant with yellow/green triangles and ITS-A/B AB-mix with green rectangles.

4.4. Discussion

The present study on the population structure of *Pseudo-nitzschia multistriata* benefits from the large number of strains obtained over a series of sample dates in the fall of 2008 and the summer and fall of 2009 because it permits statistical testing of statements inferred from the data. Moreover, this study benefits from the multiple types of data gathered from these strains (*e.g.* ITS-types and cell size) and from the insights obtained from earlier studies (2004-2006, D'Alelio *et al.*, 2009a) on this species in the Gulf of Naples. The cell size, the sample season, the potential cohort and the ITS-type was monitored for each strain collected. The ITS patterns permit

rudimentary population genetic inferences. Measurements of cell lengths, conducted immediately after strain establishment, permit tentative assignment of strains to cohorts in the two-year life cycle postulated by D'Alelio *et al.* (2010) (*i.e.* with sexual reproduction occurring in 2005, two cohorts should be present in 2006). Comparisions were conducted also with ITS patterns between years, between sample seasons within years, and potentially between cohorts.

Knowledge of the life cycle of *P. multistriata* and of the timing of the sexual phase is essential for understanding population genetic patterns and changes therein. The present data on length distribution of the sampled cells of *P. multistriata* are in agreement with the two-year life cycle proposed by D'Alelio *et al.* (2010). According to predictions by D'Alelio and co-workers, 2007 was a year in which sexual reproduction occurred. The year 2008 should then show a 2007-cohort of tall cells and its parental 2005-cohort of smallish cells, not longer sexually active. Indeed, in the 2008-samples there were a few tall cells and numerous small ones. Instead, 2009 shows only a single size class whose average cell length has diminished from the summer samples to the autumn samples. According to the model proposed by D'Alelio and co-workers the 2009-samples are obtained exclusively from the 2007-cohort, which has become sexually mature in 2009. The small cells in the 2008 sample represent the parental cohort of the cells sampled in 2009. That is, of course, if there is only one resident population of *P. multistriata* in the Gulf of Naples or if there is only one resident population in all of the Mediterranean Sea. The generated dataset does not permit testing if the parental 2005-cohort sampled in 2008 was genetically different from the 2007-cohort because the number of members of the 2007-cohort in the 2008 sample was too small.

The present data do not support the hypothesis of a merger of a *P. multistriata* population with ITS A-type with one of a ITS B-type in the Gulf of Naples. If the small cell-size group of 2008 sample is representative of the F1 2005-cohort sampled in 2006, and then the 2008 sample should contain hardly any strains exhibiting the pure-B ITS-type. Instead, consequent strains in the 2008 sample exhibit ITS B-type. Neither should the 2008 sample contain a massive proportion of strains of the ITS-A/B B-dominant type because the proportions of strains of ITS-B and ITS-A/B B-dominant type in the 2006 sample were low. Also unexpected are the marked increase of the proportion of pure ITS-types from 2008 to 2009 and the shift from a large proportion of strains of ITS-A/B B-dominant in 2008 to a large proportion of strains of ITS-A/B A-dominant in 2009. Thus, it is unlikely that the 2006 population would generate (in 2007) an F1 population exhibiting a distribution of ITS-types as deviant from it as shown in the 2008 population. Likewise, the tall cell of 2008 can in no way generate a 2009 population with such a different composition of ITS-types.

In theory, pairs of ITS-A/B A-dominant parents could produce offspring with a pure ITS-A. Likewise, pairs of ITS-A/B B-dominant parents could produce offspring with a pure ITS-B. That line of reasoning can explain the proportions of ITS-types in the 2007-cohort (small cells sampled in 2009), as a result of sexual reproduction from ITS-types proportion in the 2005-cohort (smaller cells sampled in 2008). Actually, smaller cells in 2008 (2005-cohort) were ITS-A/B A-dominant and ITS-A/B B-dominant; a sexual event between two A-dominant, two B-dominant or one of each kind may give rise to respectively ITS-A type and ITS-A/B A-dominant, ITS-B type and ITS-A/B B-dominant and finally ITS-A/B AB-mix in 2009 as barely seen in the results.

To recover a pure ITS-type from a merge ITS-A/B, the number of rRNA cistrons per haploid genome may be extremely low in *P. multistriata.* In that case, parental cells of an ITS A/B-type could produce gametes of a pure ITS A-type or B-type and thus, generate F1 cells of a pure ITS A-type or B-type. However, such a low number of cistrons is unlikely for two reasons: First, cistron copy number is correlated to cell size ((Creach *et al.*, 2006; Zhu *et al.*, 2005); see Fig. 4.1), putting an estimate of the copy number of *P. multistriata* roughly around 100 copies; Second, in an ITS-cloning exercise of the PCR-products generated from an ITS A/B-type strain D'Alelio *et al.* (2009a) observed several types of recombinants between the ITS A-type and B-type among a majority of clones of the ITS A-type and ITS B-type.

The results show no evidence for the possibility that different sub-populations occur in different seasons because the proportions of the different ITS-types were similar in the summer and autumn collections of strains in 2009. Apparently, the population(s) encountered in the summer in the Gulf of Naples is the same as that encountered there in the fall.

Concerted evolution by mean of gene conversion and unequal crossing-over cannot explain the observed results. Concerted evolution results in homogenization of the ITS variation, with one of the various ITS-types becoming dominant. Concerted evolution does not explain the large proportion of ITS A-type and ITS B-type cells generated from a parental cohort with a much smaller proportion of cells with those pure ITS A- or B-types. It cannot either explain the shift of dominant ITS-A/B type every two years (*e.g.* ITS-A/B A-dominant in 2006, ITS-A/B B-dominant in 2008, and ITS-AB no-dominance in 2005 and 2009). However, an interesting correlation is that every year of sexual reproduction there is no clear dominance of an ITS A/B-type

whereas in every year of vegetative phase monitored a clear ITS-A/B dominance is seen with the potential recovery of a ITS-pure type.

The easier *ad hoc* hypothesis is that two sub-populations co-occur in the Gulf of Naples, each one belonging to a cohort, as validated by cell size pattern. One sub-population may be composed of parental strains and the other of F1 strains, that are genetically different but correlated. The 2008 sample may have included cells from the two sub-populations, one with a distribution of ITS-types similar to that of the one sampled also in 2009 (2007-cohort), and the other one of parental strain (2005-cohort). This theory would explain the high proportion of ITS-B recorded in 2009, as a result of sexual reproduction between parental strains (ITS-A/B B-dominant) from 2007-cohort recorded in 2008 samples.

Another possibility is that the two sub-populations are no longer divided in cohorts but by ITS-types. ITS A-type and ITS-A/B A-dominant cells might belong to a sub-population whose timing of sexual reproduction is slightly offset from that of a sub-population consisting of cells of the ITS B-type and the ITS-A/B B-dominant type, at least in the field. The marked increase in the proportion of the ITS A-types in the 2009 samples could then result from particularly successful sexual reproduction of the A-type population in 2007. If this is correct then a population genetic survey using microsatellite markers should reveal two sub-populations correlated with the ITS patterns, existing side by side in the Gulf of Naples.

A third possible explanation for the radically shifting proportions of the various ITS-types over the years is that each year the Gulf of Naples contains a different population or a different set of populations. A glance at the distribution of ITS-types over the numbers of strains sampled in 2008 and 2009 shows that the numbers of strains of a particular ITS-type change radically for the number of strains with the

ITS-A/B B-dominant pattern, which is vastly higher in 2008 than in 2009. ITS-B fluctuations and the un-expected recovery of ITS-A in 2009 may be explained *ad hoc* by the existence of two populations: one of ITS-B and ITS-A/B B-dominant and the other consisting mainly of ITS-A and ITS-A/B A-dominant, which is not longer present in the Gulf in 2008; or simply not abundant enough to be sampled in 2008. If this hypothesis is correct then populations may circulate freely into/out of the GoN in currents. The population genetic survey cited above should reveal two populations in the Gulf of Naples, as a shift every year of one main population in 2008 and two populations in 2009 samples.

However a fourth hypothesis may suggest that there is no link between ITS-types, cohorts and population structure. The Gulf of Naples is influenced by currents with potential shifts of resident populations that separately undergo sexual and vegetative reproductions but may be in synchrony with environmental conditions. For example one population may got installed before 2005 (population with a high content of ITS A-type), merged with the previous resilient population, and progressively began to flow away from the gulf. In 2007 another population (mainly composed of ITS B-types) enter the GoN, merged and flow and so on. This hypothesis may explain the alternative pattern seen every odd year in the GoN. The population genetic survey should reveal, as cited above, two populations in the Gulf of Naples, as a shift every year composed of one main population in 2008, two populations in 2009 samples and intermediate strains, resulting of population merging but not part of them.

Results of a population genetic study by Evans *et al.* (2005) revealed that *Pseudo-nitzschia pungens* in the North Sea formed just a single population. The North Sea constitutes a shallow, eutrophic epicontinental sea about half the size of

the western Mediterranean, and with strong current patterns, driven by wind, tides and Atlantic currents. *Pseudo-nitzschia pungens* blooms everywhere in it. Instead, the western Mediterranean includes several deep basins with oligotrophic water, fringed by narrow, shallow, mesotrophic and locally eutrophic coastal pockets in which coastal phytoplankton can blooms year-round. Appearances of *P. multistriata* are probably restricted to the patchy coastal pockets. Current patterns in the Tyrrhenian Sea are such that local populations are partly flushed out into the open oligotrophic sea, but strong, near-shore, long-distance currents among the coastal pockets are rare. This would allow for regional populations within a Mediterranean metapopulation, each of these persisting temporarily in coastal pockets. Differences in proportions of ITS-types between years could result from incidental flushing of the coastal pockets, partially mixing the populations among pockets or even replacing a resident with another one. More refined population genetic markers, such as microsatellites are needed to screen the samples for evidence of such changes of populations inhabiting the Gulf of Naples.

Last but not least, the number of non-read strains (169 strains over 2008 and 2009) has to be considered. The proportion of reading failures in the 2009 samples was much higher than those in the 2008 samples. The failures corresponded to strains for which the ITS-sequence were unreadable because of a rapid decrease in signal strength of the electropherogram, low DNA concentration, technical difficulties of the sequencing platform, incorrectly amplified PCR-products, or difficulties with the purification of the PCR-products. One difference between the treatment of the 2008 and 2009-samples was that the latter were PCR-re-amplified predominantly at the SZN Molecular Biology Service (SBM). This can impact on the DNA quantity in the product and therefore decrease the signal intensity.

CHAPTER FIVE

DEVELOPMENT AND USE OF MICROSATELLITE MARKERS IN THE PLANKTONIC DIATOM *PSEUDO-NITZSCHIA MULTISTRIATA*

5.1. Introduction

Since 1995 the pennate diatom *Pseudo-nitzschia multistriata* is a member of the plankton flora in the Gulf of Naples (Ribera d'Alcalà *et al.*, 2004; Sarno & Dahlman, 2000; Zingone *et al.*, 2006). The species is recorded in plankton samples at the Long-Term Ecological Research station MareChiara (LTER-MC) from the early summer into late autumn. D'Alelio *et al.* (2009a) observed marked genetic differentiation in the Internal Transcribed Spacer (ITS) region of the nuclear rDNA sequences among *P. multistriata* strains collected in the Gulf of Naples. Yet, these differences did not affect the sexual compatibility between *P. multistriata* strains exhibiting distinct ITS-types and neither did they affect the viability of the offspring of such crosses (D'Alelio *et al.* 2009a). To better understand the nature and extent of this peculiar ITS variation in the population of *P. multistriata* in the Gulf of Naples, a series of polymorphic nuclear microsatellite markers has been developed for this species for screening purposes.

Microsatellites are highly polymorphic as a result of their peculiar structure. The core regions of these markers are composed of short, tandemly repeated sequences of two to six-nucleotides (Fig. 5.1). The number of repeats changes frequently because of a high propensity of insertions/deletions of single repeated units resulting from enzyme slippage during mitotic DNA replication or from unequal crossing-over during meiosis (Selkoe & Toonen, 2006). Microsatellite mutation rates

have been estimated *in vivo* to be as high as 10^{-2} to 10^{-5} events per locus per replication in bacteria (*Escherichia coli*) and yeast, respectively ((Henderson & Petes, 1992; Levinson & Gutman, 1987); (Strand *et al.*, 1993) cited in (Hancock, 1999)). Apart from being highly polymorphic, microsatellites are co-dominant, supposedly neutral and inherited in a Mendelian fashion. The markers confer a high level of accuracy in individual-specific genotyping, and the generated data can be evaluated using rigorous statistical procedures to test population genetic theories (Adams *et al.*, 2009; Evans & Hayes, 2004; Evans *et al.*, 2005; Rynearson & Armbrust, 2000). For all these reasons, microsatellites have been the markers of choice for population genetic studies.

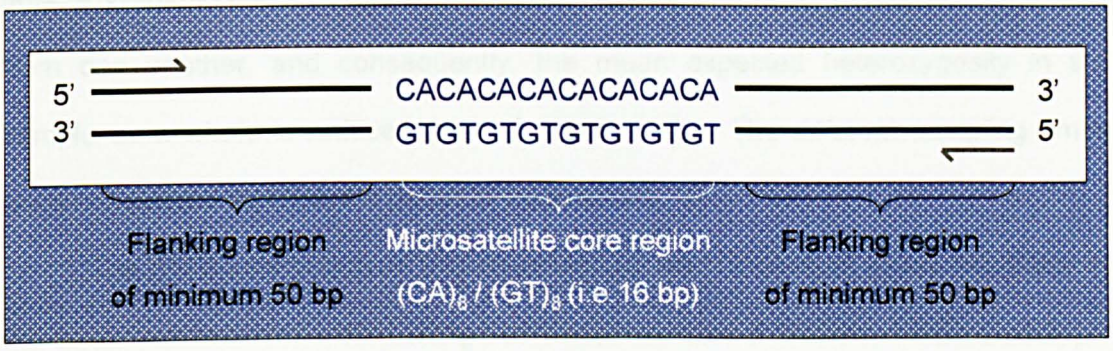


Figure 5.1: Microsatellite sequence cartoon (i.e. PNm7); DNA flanking region in black bold; reading frame from 5' to 3' (gray arrow); microsatellite core repeat (CA)8 / (GT)8 (i.e. 16 bp total); minimum distance between flanking region and microsatellite core of 50 bp.

Seven polymorphic microsatellites were characterized for *P. multistriata* in the Gulf of Naples. Polymerase chain reaction (PCR) primers and conditions for multiplex amplification were optimized for this experiment. Assessment of microsatellite polymorphism was performed on strains collected in four plankton net samples taken at different days during the autumnal bloom of 2008. Applicability of *P. multistriata* microsatellite primers on this species beyond the Gulf has been tested using co-specific strains from Portugal and Spain. Primer cross reactivity has been tested

against two other species, namely the congeneric species *Pseudo-nitzschia pseudodelicatissima* (Hasle) Hasle and the distantly related centric diatom *Leptocylindrus minimum* Gran.

5.2. Materials and Methods

Strain isolation and microsatellites detection and analysis were described in Chapter 2. The number of strains analysed was 60, taken from four sampling times (15 strains per sampling point) during the 2008 autumnal bloom. To avoid any Wahlund effect, the seven microsatellites were tested on 15 strains per isolation-time. A Wahlund effect occurs when the numbers of strains per group differ markedly from one another, and consequently, the mean expected heterozygosity in the sample as a whole is reduced (Hartl & Clark, 1989). The different sampling times were considered as separate potential populations.

The multilocus genotype was assessed for each individual strain. Neutrality of the seven markers was tested using LOSITAN (Beaumont & Nichols, 1996). Genotypic diversity within sampling points and within the whole data set was estimated with GIMLET (version 1.3.3, (Valière, 2002)). Basic population genetic parameters (number of alleles (Na), number of private alleles (*i.e.*, alleles only present in a single population), population subdivision (F_{ST} ; *i.e.*, population genetic variance among the groups), inbreeding coefficient (F_{IS}), number of migrants (Nm), Nei genetic distance and Hardy-Weinberg equilibrium (HWE)) were assessed with GENALEX (version 6, (Peakall & Smouse, 2006)). The expected heterozygosity (H_{exp}) corresponded to the mean heterozygosity of populations as a whole weighted by the number of strains tested. When the observed heterozygosity (H_{obs}) is higher than the expected heterozygosity ($H_{obs} > H_{exp}$), the population presents an excess of heterozygotes;

while a lower H_{obs} than H_{exp} indicates an excess of homozygotes. Linkage dis-equilibrium was calculated with GENETIX (version 4.02), while evidence of null alleles was shown with MICROCHECKER (version 2.2.3, (Shipley, 2003)) and population structure was assessed thanks to STRUCTURE (version 2.1, (Pritchard et al., 2000)).

5.3. Results

Seven species-specific microsatellite markers (*PNm1, PNm2, PNm3, PNm5, PNm6, PNm7* and *PNm16)* have been isolated from *P. multistriata*. Their PCR-amplification and fragment analysis protocols were optimised (refer to Chapter 2). Positive amplification of microsatellite markers was observed for all *P. multistriata* (Gulf of Naples, Spain and Portugal) while neither *P. pseudodelicatissima* nor *Leptocylindrus mimimum* gave PCR amplification products (Table 5.1).

Microsatellites were neutral candidates (Fig. 5.2), meaning that their expression was not biased by natural selection or organism fitness but constitute good tools to identify genetically conserved groups of organism. All loci were polymorphic, with the number of alleles per locus ranging from 3 to 10 (Table 5.2). Forty-four distinct genotypes were identified among 60 strains analysed, indicating a relatively high genetic diversity of 73.33%.

Observed heterozygosity (H_{obs}) ranged from 0.250 to 0.967 while expected heterozygosity (H_{exp}) ranged from 0.337 to 0.720 (Table 5.2). Observed heterozygosity was higher than the expected heterozygosity, except for loci *PNm1* and *PNm2,* which showed an excess of homozygotes. The negative mean F_{IS} (-0.310) confirmed the overall excess of heterozygosity. Nevertheless, no significant deviation was observed from HWE ($p < 0.05$) over the sampling period and no linkage dis-equilibrium was found over 1% of error.

Little structure and low divergence were found among the four groups of strains (F_{ST} = 0.021). Actually, they were also characterised by low divergence among groups (AMOVA = 2%), small Nei distances [0.015-0.051], a high number of migrants (mean Nm = 11.886) and a low number of private alleles [0-0.571] per group.

Table 5.1: Test of species-specific amplification of microsatellite region, designed for Pseudo-nitzschia multistriata *species,* on P. multistriata *(positive test: Spain and Portugal) and* P. pseudodelicatissima *and* Leptocylindrus minimum *(negative test: Gulf of Naples, Italy). The test was assessed on the seven microsatellite developed (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1) over the two alleles per locus. A "-" indicates missing i.e., no microsatellite amplification for that locus.*

Species	Country	*PNm2*	*PNm5*	*PNm7*	*PNm3*	*PNm6*	*PNm16*	*PNm1*
Pseudo-nitzschia multistriata	Portugal	175 177	238 240	262 262	– -	– -	– -	117 117
Pseudo-nitzschia multistriata	Spain	175 175	236 238	258 262	202 205	247 265	325 331	117 177
Pseudo-nitzschia pseudodelicatissima	Italy	– -	– -	– -	– -	– -	– -	– -
Leptocylindrus minimum	Italy	– -	– -	– -	– -	– -	– -	– -

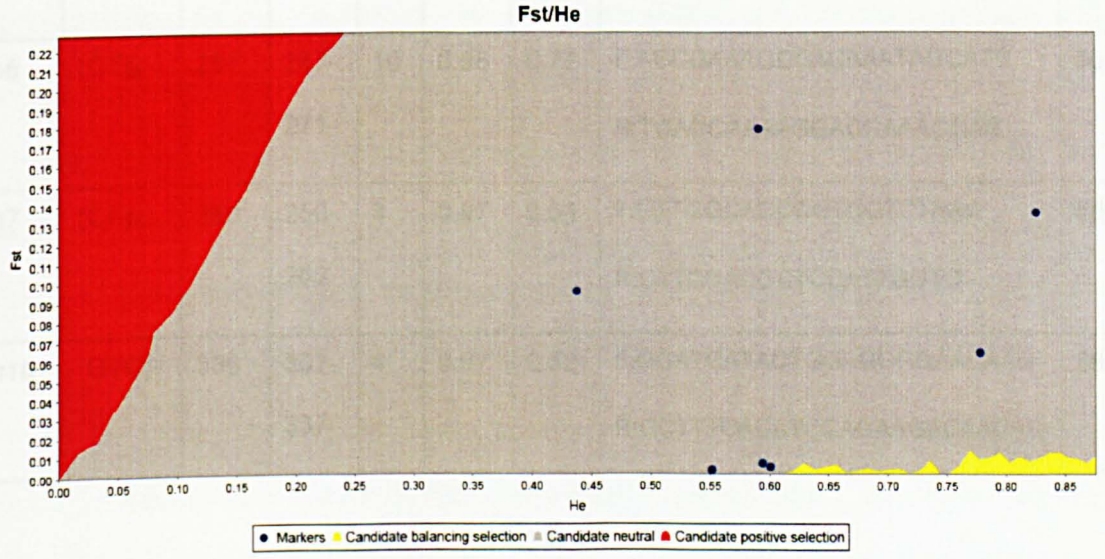


Figure 5.2: Test of neutrality of the set of microsatellites developed for the species Pseudo-nitzschia multistriata *(PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1) using* LOSITAN.

Table 5.2: Characteristics of the seven microsatellite primers: using GENALEX. Core denotes core motif, Size indicates size of the original sequence, SR stands for size range of the alleles encountered, Na denotes number of alleles per locus, Hobs and Hexp signify observed and expected heterozygosity, Primer sequences: F is forward primer, R is reverse primer, T is the primer annealing temperature in °C.

Locus	Core	Size	SR	Na	H_{obs}	H_{exp}	Primer 5' to 3' sequence	T °C
PNm1	$(TC)_{23}$	149	115-147	8	0.28	0.54	F:CACCAATTGCATCCTAAAAGGG R:TCCGTCTAAGCCTGTATTTGTGAC	54
PNm2	$(AC)_{17}$	189	175-187	5	0.25	0.34	F:GGGATCGATTCGTGAAAGAGC R:GCATAGAAGCACGGCACAGTG	58
PNm3	$(GAC)_8$	208	202-208	3	0.87	0.50	F:GGATCGAATAGGGGATGAATACG R:GGAGCTTGCATCATCATCACAG	58
PNm5	$(GT)_{11}$	240	224-244	6	0.73	0.70	F:GAACAGAACTGCCCGAAGGAC R:AGGATCACCCACGAGACACTG	58
PNm6	$(CT)_9$	251	247-271	10	0.98	0.72	F:AGCGAAAGCGACAAATAGCATC R:TGAGCAAAAGGACGAAACGAG	58
PNm7	$(CA)_8$	260	250-262	3	0.97	0.51	F:GTTGGCACCGGTGGTCTAAC R:CTTCGACGCTCCATTGGTG	58
PNm16	$(GTC)_7$	336	307-337	4	0.97	0.52	F:GGATCATACTGGAGGGGAACAAG R:GCTTTCACATCCAGAAGACAACAG	58

5.4. Discussion

The seven microsatellite markers identified in this study are species-specific and polymorphic. No clear structure was found among the 60 strains over the four sample dates of the autumnal bloom of 2008. The strains may belong to one single population or to two or more within a meta-population structure. To choose between these possibilities, genetic analyses were conducted on a much larger set of *P. multistriata* strains collected over three blooming periods (in 2008 and 2009; see Chapter 6). *Pseudo-nitzschia multistriata* grows clonally during most of its life cycle, but the genetic diversity encountered among the strains in the samples indicates the occurrence of sexual reproduction.

The slight but insignificant deviation from Hardly-Weinberg equilibrium ($p < 0.05$) for each of the loci tested may suggest population subdivision, sex-linkage, or overlapping between generations. The slight, but insignificant, excess of heterozygotes on the loci *PNm5, PNm7, PNm3, PNm6* and *PNm16* could hint at a bias by ecological and behaviour processes. It cannot be due to the so-called Wahlund effect because sample size was equal for each of the sub-samples. The slight but insignificant excess of homozygotes on *PNm1* and *PNm2* could be due to the presence of null allele at these loci, to errors of alleles scoring, to clonal reproduction of the species and/or to effects related to small samples.

Null alleles occur when a base change in a primer region of a particular allele leads to amplification failure of that allele (Callen *et al.*, 1993; Estoup & Angers, 1998; Oetjen & Reusch, 2007). If a locus is homozygous for that allele, the microsatellite will not be amplified, and the locus is scored as 'missing data' (Selkoe

& Toonen, 2006). If the locus is heterozygous (*i.e.* one allele with- and the other without base change in its primer region) then only the allele without base change will be amplified and the locus will be scored as homozygous. Null alleles lead to mis-interpretation of data. To identify null alleles, crossing experiments have been conducted and new primers have been designed. Newly designed primers were aspecific and the crossing experiments did not reveal the presence of null alleles on the different loci tested (Chapter 7).

Homozygote excess may be linked also to allele scoring errors such as (i) PCR amplification errors and (ii) pherogram analysis mistakes. PCR errors may occur when the quantity of DNA to replicate is too low (Gagneux *et al.*, 1997) or when one allele is preferentially amplified (*e.g.*, because of a much smaller length (Wattier *et al.*, 1998)). Consequently, one of the two alleles is amplified in a higher quantity, lowering the relative signal intensity of the other allele. Analysis errors of microsatellite electropherograms can happen in case of high signal intensities of two alleles with similar lengths (*e.g.*, only one duplet difference). In that case the two peaks will overlap, appearing as one peak and masking the heterozygosity. To minimize such scoring errors, pherograms that showed mis-amplification, intense signals or unclear class categories were re-amplified or reloaded using a lower quantity of the amplification product.

Homozygote excess can be a consequence of biological, meterological and physical processes in combination with low sampling effort. Yet these are not issues associated to the microsatellite markers, but issues that arise because of sampling strategies, evaluation of results and biological peculiarities of the species studied. Such issues will be addressed as they might arise in the much more extensive studies of *P. multistriata* presented in Chapters 6, 7 and 8.

5.5. Conclusions

The seven microsatellite markers selected for *Pseudo-nitzschia multistriata* can be applied to study population genetic patterns and processes within and among populations of this species in the Gulf of Naples.

CHAPTER SIX

MICROSATELLITE DIVERSITY ACROSS BLOOMS OF *PSEUDO-NITZSCHIA MULTISTRIATA*

6.1. Introduction

In the present Chapter intra- and inter-population genetic patterning is assessed in the planktonic pennate diatom *Pseudo-nitzschia multistriata* in the Gulf of Naples (GoN). There, the species was observed for the first time in 1995 in the plankton counts at the Long-Term Ecological Research station MareChiara (LTER-MC), and it has appeared in the autumn of the years since then. From 2004 onwards it blooms also briefly in the summer (June - August). The species exhibits a biennial life cycle with a sexual reproduction occurring only in odd years (D'Alelio et al., 2010). Initial cells are about 80 μm in length, much taller than their parents, but with ongoing mitotic divisions the cells become progressively smaller, and after two years have reached a size window in which cells can become sexualized. Those that fail to reproduce sexually keep on dividing mitotically into their third year and then perish. Cohorts can be identified by cell size distribution. In years in which sexual reproduction occurs, the size class of cells of the two-years old parental cohort dominates; the initial cells of their offspring cohort are rare. The next year shows two size classes of cells: the now one-year old offspring cohort of relatively tall cells, and the now three-years old remainder of the parental cohort consisting of minute cells.

Strains of *P. multistriata* sampled in the GoN between 2004 and 2006 also showed remarkable genetic differences on the internal transcribed spacer regions (ITS-1 and ITS-2) flanking the ribosomal RNA coding regions in the rDNA cistrons

(D'Alelio et al., 2009a). D'Alelio *et al.* (2009a) exhibited a pure ITS-type (ITS A-type of ITS B-type) and a mixed type (ITS A/B-type). Strains interbred successfully irrespective of their ITS-type, and the progeny was also fertile. Therefore, D'Alelio *et al.* (2009a) concluded that in the GoN *P. multistriata* constitutes a single biological species.

Results of a recent study (Chapter 4) assessing the ITS-types of a far larger number of *P. multistriata* strains sampled in the GoN in 2008 and 2009 showed that the 2005-cohort (sampled in 2008) could not have given rise in its entirety to the 2007-cohort (sampled in 2009) given the radically different distributions of the proportions of their ITS-types. These results suggest the existence of multiple sub-populations of *P. multistriata* in the GoN, each with distinct proportions of ITS-types. Differential sexual reproduction success of these sub-populations might explain the radical shifts in proportions of ITS-types between one cohort and the next. Alternatively, different sets of sub-populations may inhabit the GoN in different years, implying that *P. multistriata* in the Mediterranean Sea is composed of several populations scattered along coastal pockets. Similar proportions of strains of the recognized ITS-types in the summer and autumn samples of 2009 suggest that the population(s) remain in the GoN at least throughout the growing season.

The present study takes the results obtained by the ITS studies (Chapter4) as a starting point to assess population genetic structure and dynamics in *P. multistriata* in the GoN using seven microsatellite markers (Chapters 2 and 5). The survey includes strains utilised in the ITS-studies to permit comparison of the ITS- and microsatellite-data. If sub-populations are detected, then sample seasons, cohorts and ITS-types will be mapped over the population assignment based on microsatellite data to assess if relationships exist.

Microsatellites are ideal markers to address population genetic structure. These markers are DNA sequences whose core regions consist of short, tandem repeats of two to six-nucleotides (Hancock, 1999). The number of repeats varies among individuals because enzyme slippage during DNA replication and unequal crossing-over during meiosis result in insertions/deletions of single repeated units. Therefore, these markers exhibit high mutation rates and high polymorphism levels (*i.e.* a high number of allele versions). Moreover, they are selectively neutral, species-specific, co-dominant, inherited in a Mendelian fashion and independently from one another (Selkoe & Toonen, 2006). For those reasons, microsatellites are used for assessing intra-specific variability, migration rates among populations, population size, kinship (paternity, clonal structure) and the occurrence of past bottlenecks (Selkoe & Toonen, 2006). Multiple microsatellite markers can be PCR-amplified in one and the same reaction mix (multiplex) to augment the number of microsatellites amplified per PCR reaction.

There are of course, also some downsides to the use of microsatellites. First they are expensive to develop and relatively laborious to score. Moreover, missing values and null alleles, Wahlund effect, linkage dis-equilibrium and homoplasy need to be taken into account during data analysis (Selkoe & Toonen, 2006). Null alleles correspond to the non-amplification of an allele because of a mutation in one of its primer sites (Callen *et al.*, 1993; Oetjen & Reusch, 2007). Consequently, heterozygous specimens are mis-identified as homozygous. The Wahlund effect refers to reduction of heterozygosity caused by subpopulation structure (Hartl & Clark, 1989). If the subpopulations have different allele frequencies, then the overall heterozygosity is reduced, even if the subpopulations themselves are in Hardy-Weinberg equilibrium (HWE). Linkage dis-equilibrium occurs when the observed

combination of two alleles differs markedly from the expected frequency based on a random haplotype expression. It occurs with alleles located either proximal on the same chromosome or with allele combination that are functionally linked (Devlin & Risch, 1995; Reich & Goldstein, 1999). Homoplasy occurs when microsatellite alleles have the same length for other reasons than common descent. Homoplasy is not unlikely in microsatellites given their high mutation frequencies and the length of microsatellite core (Estoup *et al.*, 2002).

Unfortunately, microsatellite studies in diatoms are still rare. Rynearson and co-workers (Rynearson & Armbrust, 2000; Rynearson & Armbrust, 2005; Rynearson & Armbrust, 2004; Rynearson *et al.*, 2009) used 2-3 markers to address population structure in the planktonic marine species *Ditylum brightwellii* (West) Grunow. The authors found population sub-division in different water masses in the Juan de Fuca Strait and Puget Sound, probably caused by adaptation to the properties of different stable water bodies over the period of the bloom. Evans *et al.* (2004) used nine loci to assess differentiation among strains of *Pseudo-nitzschia multiseries* (Hasle) Hasle from Canada, Europe, and Russia. The authors demonstrated a high differentiation among isolates from the different regions, suggesting that geographic distance fosters genetic isolation. In addition, the authors also observed differentiation among 22 strains sampled from different sites along the east coast of Canada, indicating that different populations may exist even on a regional scale. Population genetic studies in *Pseudo-nitzschia pungens* (Grunow ex Cleve) Hasle also revealed remarkable differentiation among regions. Evans *et al.* (2005) using six loci, found little spatial and temporal variation among samples of strains in the North Sea, suggestive for a single population. In contrast, Adams *et al.* (2009), using four of the six loci in Evans *et al.* (2005) observed two markedly genetically distinct populations of *P. pungens* in

the Pacific Ocean off Vancouver Island and Washington State. Each of these was genetically as divergent from the North Sea population as they were from one another. Population genetic studies of freshwater diatoms are even scarcer. Evans *et al.* (2009) deployed ten microsatellites and ITS data to assess population structure in the Freshwater benthic species *Sellaphora capitata* Mann & McDonald. The 70 isolates grouped according to their geographic origin and showed high F_{ST} values among sites (>0.4). In spite of this genetic differentiation, geographic isolates were still able to mate successfully. Nevertheless, the remarkable levels of intraspecific differentiation shown in most of the above-mentioned studies suggest limited dispersal and opportunities for allopatric speciation, which may help to explain the huge diversity of diatoms.

6.2. Materials and Methods

Of the total of 467 strains of *Pseudo-nitzschia multistriata* collected from MareChiara net samples, 412 were successfully analysed using microsatellite markers. Strains have been isolated from cells or chains of cells from the autumnal bloom 2008 (N=200), from the early-summer bloom 2009 (N=81) and the autumnal bloom 2009 (N=131) (Appendix 1). Strain culture and DNA extraction were as described in Chapter 2. Strain size was measured using light microscopy (200x magnification; Leica Microsystems DMIL, Wetzlar, Germany) after about one week of growth in culture.

Seven polymorphic microsatellites (Chapter 5): *PNm2*, *PNm3*, *PNm5*, *PNm6*, *PNm7*, *PNm16* and *PNm1* were used to study population genetic structure among the sampled strains (Appendix 2). Microsatellites were amplified in monoplex or duplex for 96 strains per PCR run using labelled-primers (CY-5 and IRD-700) as

described in Chapter 2. Following amplification, the products were purified, assembled into triplexes (*i.e.* three microsatellites) and read by a fragment analyser as explained in Chapter 2. To avoid any problems of failed amplification because of overly short or long DNA sequences, primers were designed to amplify between 150 and 350 bp products (Chapter 5 Table 5.2). Microsatellites were labelled with different fluorophores (CY-5 and IRD-700) before being associated by different length categories.

CEQ™2000XL (version 4.3.9, Beckman Coulter™) was used to infer microsatellite length. The lengths were classified in length-windows determined during establishment of the protocol (see Chapter 2). MICROCHECKER (version 2.2.3, (Shipley, 2003)) was deployed to check for errors due to allele dropout (null alleles). LOSITAN (Beaumont & Nichols, 1996) was used to test for microsatellite neutrality. All seven markers were found to be neutral. STRUCTURE (version 2.1, (Pritchard et al., 2000)) was used to assess population structure over the *P. multistriata* blooms of 2008 and 2009 without any *a priori* identification of populations. STRUCTURE employs a clustering method based on a Bayesian model, using multilocus genotype data to infer population structure and assign individuals to populations conforming to HWE and linkage equilibrium. STRUCTURE was run using the admixture and the independent correlated models, 10 000 burn-in iterations, and a run length of 50 000 iterations assuming one to distinct clusters. GENALEX (version 6, (Peakall & Smouse, 2006)) was deployed to calculate descriptive population genetic parameters including the allele fixation index (F), inbreeding coefficient (F_{IS}), population subdivision (F_{ST}), Hardly Weinberg Equilibrium (HWE), number of alleles (N_A), effective number of alleles (N_E), number of private alleles ($N_{PRIVATE\ ALLELES}$; *i.e.*, alleles only present in a population), number of alleles per locus in the data set ($N_{ALLELES/LOCUS}$), per

population ($N_{ALLELES/POP}$), per microsatellite locus and per identified population. GIMLET (version 1.3.3, (Valière, 2002)) was deployed to assess genotypic diversity of populations as delineated by STRUCTURE. GENETIX (version 4.05.2) was used to generate correspondence analyses and to calculate linkage dis-equilibrium between pairs of alleles within populations as delineated by STRUCTURE. Linkage dis-equilibrium has been tested under 1000 permutations and a threshold error of 1%.

The ITS-types of the strains as identified from the electropherograms of internal transcribed spacer regions (ITS-1 and ITS-2) on the nuclear rDNA cistrons were collected from strains as described in Chapter 4. These ITS data were compared with strain assignment to populations according to STRUCTURE to evaluate deviations from expectations.

6.3. Results

Aberrant patterns were detected in 50 of the 412 strains amplified for the seven microsatellites; one to three loci per strain failed to amplify (Table 6.1). A total of 38 of these 50 strains were from the 2008-sample. Problems were encountered principally on loci *PNm5* and *PNm7* in the 2008-sample and on locus *PNm16* in the autumn sample of 2009 (Fig. 6.1). Because GIMLET and STRUCTURE, do not accept missing-values, I only included the 362 strains for which the alleles of the seven loci were identified: these include 162 strains sampled in the autumn of 2008, 78 in the early-summer of 2009 and 122 in the autumn of 2009 (Appendix 2).

Table 6.1: Sum of missing values over all strains (N=412) per microsatellite amplified (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1) and per sampled period (early-summer (☼) 2009 and autumn (🍂) 2008 and 2009). $N_{STRAINS}$ corresponds to the number of strains showing missing values in their microsatellite pattern. The sum of missing values per sample season is indicated by $N_{MISSING\ VALUES}$. The proportion of failed amplifications is indicated for each locus and sample season, and is calculated as follow: $P_{LOCUS\ MIS\text{-}AMPLIFIED} = N_{MISSING\ VALUES} / (7\ loci \times N_{STRAINS})$.

	2008 🍂	2009 ☼	2009 🍂
$N_{STRAINS}$	38	3	9
$N_{MISSING\ VALUES}$	58	5	13
$P_{LOCUS\ AMPLIFICATION\ FAILURE}$	0.22	0.24	0.21
1 $N_{LOCUS\ AMPLIFICATION\ FAILURE}$	0.3	0.2	0.5
2 $N_{LOCUS\ AMPLIFICATION\ FAILURE}$	0.6	0.8	0.3
3 $N_{LOCUS\ AMPLIFICATION\ FAILURE}$	0.1	0.0	0.2

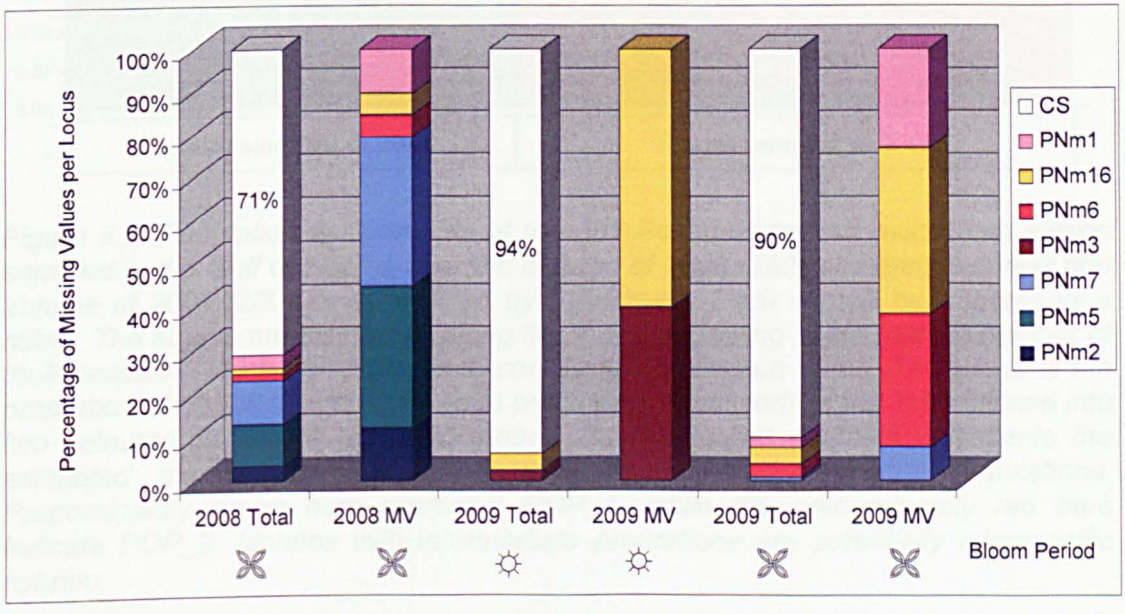


Figure 6.1: Distribution of missing data over the seven loci for each of the three sample periods tested (early-summer (☼) 2009 and autumn (🍂) 2008 and 2009) expressed in %. All data (complete sequence (CS) and those presenting missing values (MV)) for one sample period sum up to 100% (e.g. in the fifth column 90% of complete sequence and 10% of the missing data). All missing values for one sample period sum up to 100% (e.g. in the sixth column 30% of the missing values were due to no-shows on PNm6).

STRUCTURE identified two populations from here onwards denoted as POP_1 and POP_2. A total of 122 strains (marked green in Fig. 6.2) were assigned to POP_1 and 227 strains (marked red in Fig. 6.2) to POP_2. Thirteen strains could belong to either population or to a third one and will be referred to as the 'Intermediate strains' from here onwards. Both populations were present during the autumn of 2008 whereas the 2009-samples almost exclusively belonged to POP_2 (see Appendix 1 for strain assignment).

Figure 6.2: Population assignments of the 362 Pseudo-nitzschia multistriata *strains sampled in the Gulf of Naples over the autumn of 2008 (162) and early summer and autumn of 2009 (200) as determined by* STRUCTURE. *Each vertical bar represents a strain. The strains are distributed along the x-axis according to the ranking number of their isolation, which corresponds to consecutive collection dates. The y-axis is the proportion of alleles derived from each population. Each vertical bar is partitioned into two coloured segments (red and green). Each coloured segment represents the estimated membership fraction of that strain in each of the populations. Predominantly green bars represent POP_1, while the predominantly red bars indicate POP_2. Isolates with intermediate proportions are potentially interspecific hybrids.*

Table 6.2a: Allele parameters for each of the populations (POP_1, POP_2 and Intermediate strains), obtained using GENALEX: Total number of alleles per population ($N_{ALLELES\ TOTAL}$), Number of alleles per locus ($N_{ALLELES/LOCUS}$), Number of alleles per locus that are significantly present ($N_{ALLELES/LOCUS\ (p \geq 0.05)}$), Effective number of alleles per locus ($N_{EFFECTIVE\ ALLELES/LOCUS}$), Number of private alleles per locus ($N_{PRIVATE\ ALLELES}$) and Allele frequency ($N_{EFFECTIVE\ ALLELES/LOCUS} / N_{ALLELES/LOCUS}$).

	POP_1	POP_2	Intermediate strains
$N_{ALLELES\ TOTAL}$	46	57	31
$N_{ALLELES/LOCUS}$	6.6	8.1	4.4
$N_{ALLELES/LOCUS\ (p \geq 0.05)}$	2.4	3.0	2.6
$N_{EFFECTIVE\ ALLELES/LOCUS}$	2.3	2.7	2.5
$N_{PRIVATE\ ALLELES}$	1.1	3.1	0.1
Allele frequency	0.348	0.333	0.568

Table 6.2b: Allele parameters for each of the alleles, obtained using GENALEX: Number of alleles (N_A), Effective number of alleles (N_E), Number of alleles per locus in the data set ($N_{ALLELES}$) and per population ($N_{ALLELES/POP}$) assessed per microsatellite locus: PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1 (top row) per population: POP_1, POP_2 and the group of Intermediate strains.

PNm	*2*	*5*	*7*	*3*	*6*	*16*	*1*	$N_{ALLELES/POP\ (\Sigma=70)}$
$N_{A\ POP\_1}$	5	**8**	6	3	**9**	6	**9**	46
$N_{E\ POP\_1}$	1.3	4.2	2.2	2.0	2.5	2.1	1.6	16
$N_{A\ POP\_2}$	5	**10**	7	5	**14**	6	**10**	57
$N_{E\ POP\_2}$	2.0	1.8	2.1	1.4	4.7	3.3	3.3	19
$N_{A\ Intermediate\ strain}$	4	4	2	3	10	3	5	31
$N_{E\ Intermediate\ strain}$	1.9	2.4	2.0	2.1	5.3	2.2	2.0	18
$N_{ALLELES\ (\Sigma=187)}$	7	12	9	5	16	8	13	

GENEALEX uncovered high allele diversity, with 46 alleles, of which one private, in POP_1, and 57 alleles in POP_2, of which three private. However, the number of effective alleles per locus was low in each of the populations, but comparable between populations (Table 6.2a). Within POP_2 the strains collected in 2008 exhibited lower allele diversity (N_A = 37 alleles; N_E = 5 alleles) than those collected in the early-summer (N_A = 43 alleles; N_E = 9 alleles) but closeby to those collected in the autumn of 2009 (N_A = 41 alleles; N_E = 6 alleles). POP_1 and POP_2 showed similar allele frequencies whereas that of the Intermediate strains was markedly higher. Table 6.2b specifies for each population the number of alleles per locus. Numbers of alleles per locus were comparable between the two populations. Loci *PNm1*, *PNm5* and *PNm6* showed the higher number of alleles, but most among these alleles were relatively rare because effective numbers of alleles were all markedly lower.

GIMLET indicated lower genotypic diversity in POP_1 than in POP_2 and lower by far than among the Intermediate strains (Table 6.3). Thus, more similar genotypes co-existed in POP_1 than in POP_2. Within POP_2 the strains collected in 2008 exhibited a higher genotypic diversity (93.75%) than those collected in the summer and in the autumn of 2009 (88.31% viz. 58.47%). GENETIX showed no linkage *dis*-equilibrium between any pair of alleles in POP_1. However, within POP_2 linkage *dis*-equilibrium was detected between four pairs of loci involving *PNm2*, *PNm6*, *PNm7* and *PNm16* (Table 6.3).

Table 6.3: Genotype diversity (GIMLET) and Linkage dis-equilibrium (GENETIX) estimated for the two populations (POP_1, POP_2), the group of Intermediate strains (IS; estimated by STRUCTURE) and the pool data (whole data set). Genotype diversity has been calculated as follows: number of group of same genotype per number of total genotype analysed. A distinction was made between POP_2 belonging to early-summer strains (☼) and autumn strains (✂) of the year 2009.

	Genotype diversity	Linkage dis-equilibrium
POP_1	69/122 = 56.6%	None
POP_2	165/227 = 72.7% group 2008 ✂: 30/32 = 93.75% group 2009 ☼: 68/77 = 88.31% group 2009✂: 69/118 = 58.47%	*PNm2-PNm16. PNm7-PNm6. PNm7-PNm16. PNm6-PNm16*
IS	13/13 = 100%	None
All strains	247/362 = 68.2%	

Table 6.4: Population genetic parameters as calculated by GENEALEX: Fixation index (F), Inbreeding coefficient (F_{IS}), Population subdivision (F_{ST}), Hardly Weinberg Equilibrium (HWE), Number of migrants (Nm), Number of strains analysed ($N_{STRAINS}$), seven microsatellite loci: PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1, and per population: POP_1, POP_2 and the group of Intermediate strains (non assign) obtained by GENALEX. "ns" denotes HWE not significant, i.e. no equilibrium), (*) and (***) indicate significance at p<0.05 and at p<0.001, respectively. Blue marking indicate homozygote excess.

	PNm2	*PNm5*	*PNm7*	*PNm3*	*PNm6*	*PNm16*	*PNm1*	Mean
F_{POP_1}	**0.487**	-0.219	-0.769	-0.919	-0.673	-0.891	**0.402**	-0.369
F_{POP_2}	-0.432	**0.212**	**0.204**	**0.532**	-0.043	-0.272	-0.192	0.001
$F_{Intermediate}$	-0.159	**0.076**	-0.692	-0.040	-0.139	-0.724	**0.055**	-0.232
F_{IS}	-0.144	-0.016	-0.419	-0.250	-0.249	-0.593	**0.023**	-0.236
F_{ST}	0.070	0.088	0.001	0.134	0.087	0.052	0.104	0.077
Nm	3.340	2.599	171.129	1.618	2.607	4.598	2.146	3.014
	PNm2	*PNm5*	*PNm7*	*PNm3*	*PNm6*	*PNm16*	*PNm1*	$N_{STRAINS}$
HWE_{POP_1}	***	***	***	***	***	***	***	122
HWE_{POP_2}	***	***	***	***	***	***	***	227
$HWE_{Intermediate}$	***	*	*	ns	ns	*	ns	13

The inbreeding coefficient between populations, as calculated by GENEALEX, revealed an overall excess of heterozygotes (F_{IS} = -0.236). However, POP_1 showed homozygote excess on loci *PNm2* and *PNm1*, POP_2 on *PNm5*, *PNm7* and *PNm3* and the group of Intermediate strains on *PNm5* and *PNm1* (Table 6.4). If all strains were taken together, they exhibited homozygote excess only on *PNm1*. Nevertheless, all loci were in HWE within POP_1 and within POP_2, but several of the loci were not in HWE within the group of Intermediate strains. One reason maybe that the group of Intermediate strains is small, but the main reason is probably that they do not represent a single population.

Correspondence analyses of the 362 strains as determined by GENETIX showed two distinct swarms (Fig. 6.3a). Mapping of the population assignment as determined by STRUCTURE over the strains revealed that the two swarms correspond to strains assigned to POP_1 (yellow dots) and those assigned to POP_2 (white dots). Eleven of the 13 Intermediate strains (blue dots) were recovered in between the two swarms (compare Figs 6.3a and 6.3b) and possibly constitute the hybrid offspring of crosses between parents from the different populations. Two Intermediate strains (SY444 and SY483) were recovered well above the swarms. Strain SY444 exhibited a unique allele 229 on locus *PNm6* and strain SY483 showed two unique alleles (265 and 269) on the same locus. Outlying strain SY339 (Fig. 6.3b) was assigned to POP_2 and exhibited a unique allele (149) on locus *PNm1*.

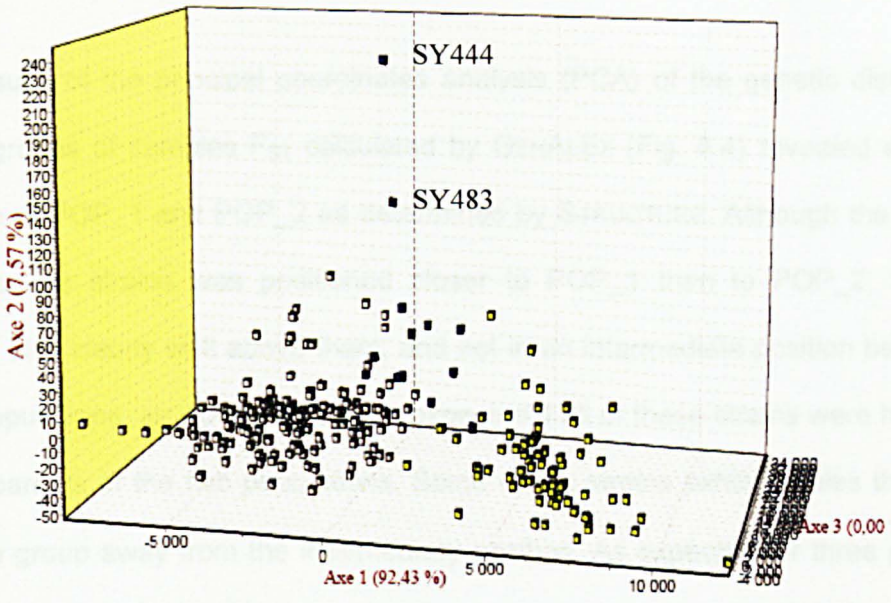


Figure 6.3a: Spatial ordination of 362 strains based on pair-wise genotypic distances among the strains, as determined by GENETIX. Yellow dots are strains assigned to POP_1, white dots to POP_2 and blue dots the group of Intermediate strains by STRUCTURE.

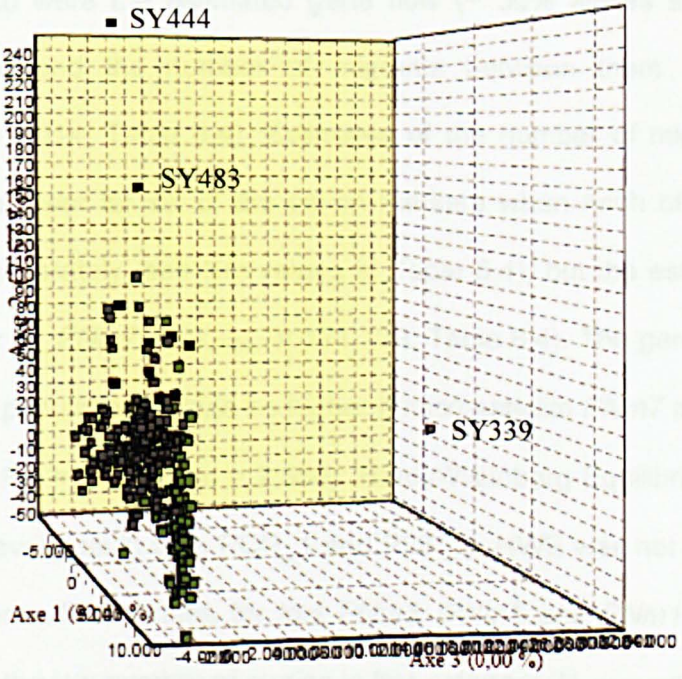


Figure 6.3b: Spatial ordination of strains as in Fig. 6.3a but viewed from a different angle to permit appreciation of 3D. Outlying strains SY444 and SY483 have been assigned by STRUCTURE to the group of Intermediate strains, outlying strain SY339 been assigned by STRUCTURE to POP_2.

Results of the principal coordinates analysis (PCA) of the genetic distances between groups of samples F_{ST} calculated by GENALEX (Fig. 6.4) revealed a clear separation of POP_1 and POP_2 as determined by STRUCTURE. Although the group of Intermediate strains was positioned closer to POP_1 than to POP_2, it was recovered also clearly well above them, and not in an intermediate position between the two populations, as would have been expected if all of these strains were hybrids between parents of the two populations. Some of the strains exhibit alleles that are pulling the group away from the intermediary position. As expected for three points, the PCA explained the full 100% of the variability of the population distribution (Axis 1: 80.16% and Axis 2: 19.84%; Fig. 6.4).

Genetic divergence between populations was low (F_{ST} mean= 0.077; Table 6.4), but so were the estimated gene flow (~ 30% alleles shared; AMOVA 999 permutations) and the number of migrants between them was small but significant (Nm = 3.014; Table 6.4). Estimates of the number of migrants between populations were lower for six of the seven markers when each of these markers were assessed separately (see Nm values in Table 6.4), but the estimated number was much higher for *PNm7* (Nm_{PNm7} = 171.129; Table 6.4). The genetic divergence between the two populations would be higher if microsatellite *PNm7* is excluded from the account (*i.e.* F_{ST} = 0.089; Nm = 2.556). Hardy-Weinberg Equilibrium (HWE) was satisfied for all seven loci in both POP_1 and POP_2. HWE was not satisfied for the group of 13 Intermediate strains for loci *PNm3*, *PNm6* and *PNm1*. The deviation could result from the low number of strains in this category ($N_{STRAINS}$ = 13; Chi-square parameter test: N>30 and Np>5; Table 6.4) or because of biological reason.

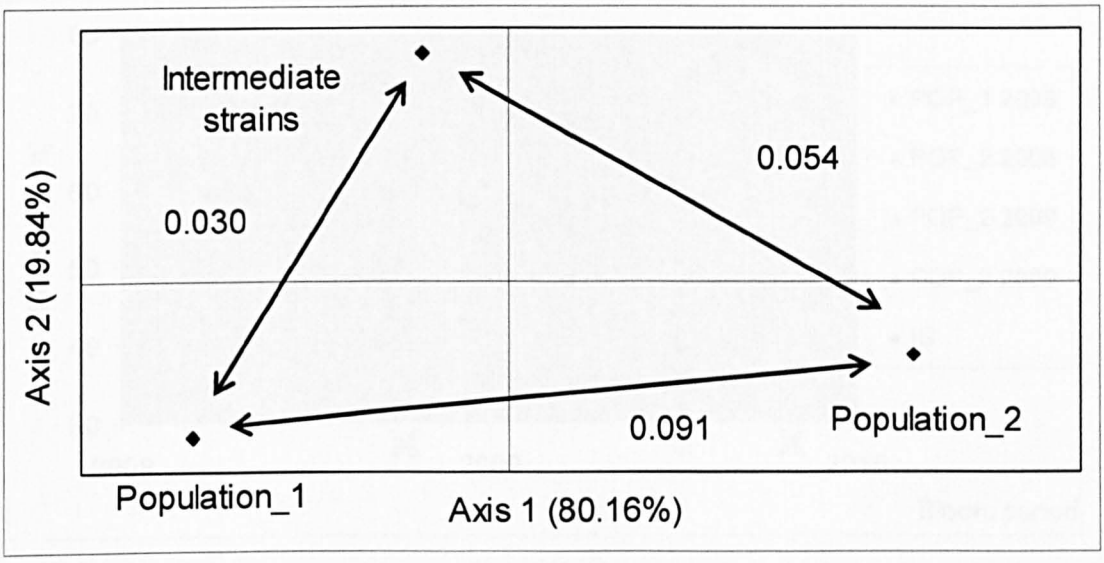


Figure 6.4: Results of Principal Coordinates Analysis (PCA) of POP_1, POP_2 and the group of Intermediate strains from F_{ST} pair-wise distance matrix (GENALEX). Black arrows and associated F_{ST} values indicate the distance between two populations.

Table 6.5: Population genotypic diversity of Pseudo-nitzschia multistriata *strains sampled during autumn 2008, early-summer 2009 and autumn 2009 ($N_{STRAINS}$=362 strains) as function of the population structure (POP_1, POP_2 and Intermediate strains) and the ITS pattern: A, B, A/B and NR (ITS pattern was not read successfully). $N_{STRAINS}$ signifies the number of strains tested per condition. G/N indicates the number of genotype groups in function of the number of strains analysed. Last row and last column indicate the diversity in the whole sample considered per population and per ITS-type.*

	POP_1		POP_2		Intermediate strains		Whole data set	
	$N_{STRAINS}$	G/N	$N_{STRAINS}$	G/N	$N_{STRAINS}$	G/N	$N_{STRAINS}$	G/N
ITS-A	0	0.0	23	95.7	1	100.0	24	95.8
ITS-B	1	100.0	46	71.7	4	100.0	51	74.5
ITS-A/B	105	56.2	52	94.2	5	100.0	162	69.8
ITS-NR	16	75.0	106	68.9	3	100.0	125	70.4
All ITS	122	56.6	227	72.7	13	100.0	362	68.2

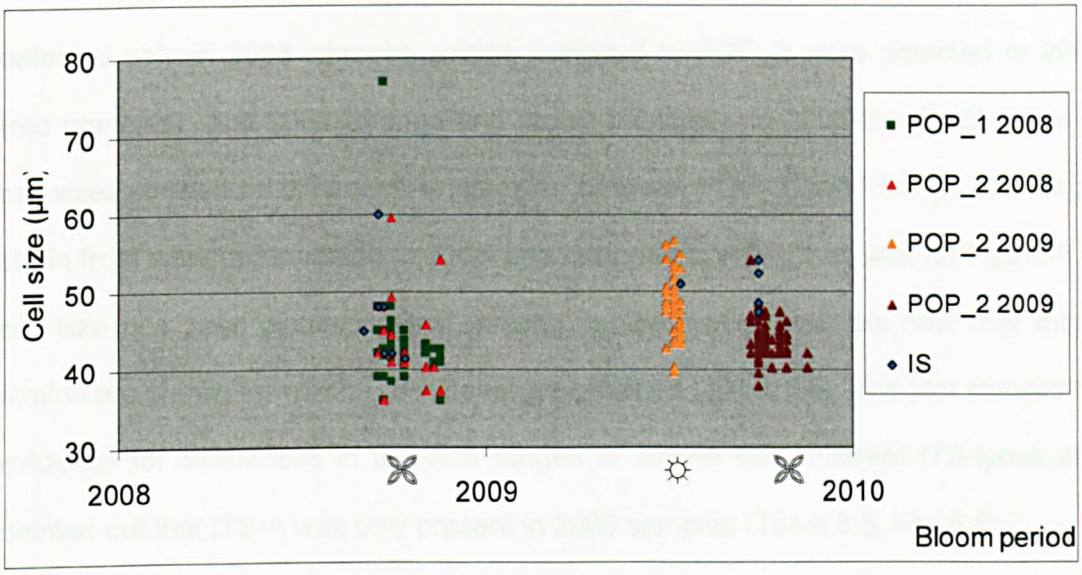


Figure 6.5a: Cell size distribution of Pseudo-nitzschia multistriata *strains (N= 362) in the populations as assigned by* STRUCTURE*; POP_1, green rectangles; POP_2, triangles (autumn (🍂) 2008, red; early-summer (☼) 2009, orange; autumn (🍂) 2009, brown). Blue diamonds denote Intermediate strains (IS).*

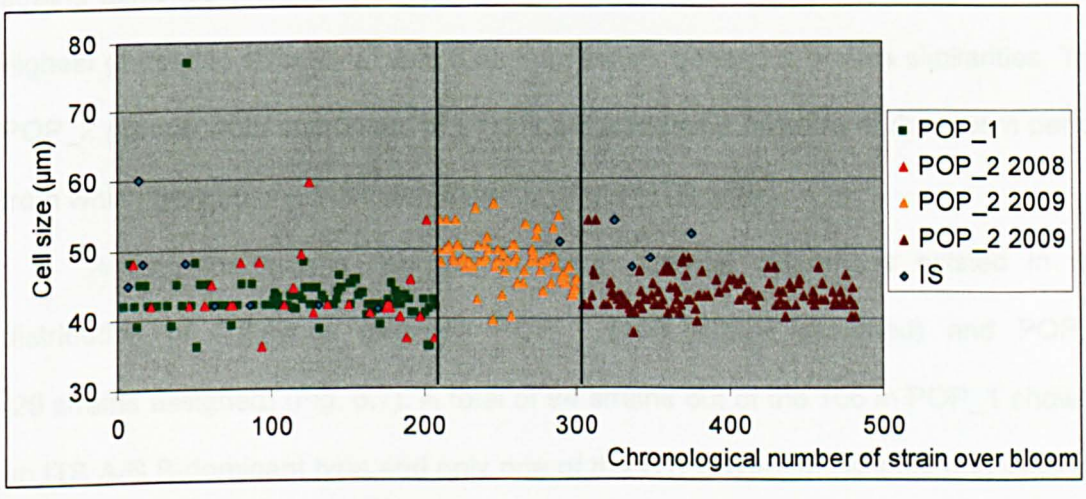


Figure 6.5b: Cell size distribution of Pseudo-nitzschia multistriata *strains (N= 362) in the populations as assigned by* STRUCTURE*; POP_1, green rectangles; POP_2, triangles (autumn 2008, red; early-summer 2009, orange; autumn 2009, brown). Blue diamonds denote Intermediate strains (IS). Cell size has been plotted against rank numbers of strains. Rank numbers reflect collection date because strains with higher rank numbers have been collected later; plotting strains according to collection date would lead to crowding of points and loss of clarity.*

In Figure 6.5 (a and b), strains assigned to POP_1 (green squares) were detected only in 2008 whereas strains assigned to POP_2 were detected in 2008 (red triangles) and 2009 (orange and brown triangles). In 2008 the distributions of cell sizes showed no difference whatsoever between POP_1 and POP_2. The single strain from a tall cell sampled in 2008 was assigned to POP_1 as well. In Figure 6.6, cell size has been plotted against ranking numbers of strains, but now only those strains are shown for which ITS-types were obtained (Table 6.5). The plot showed no evidence for differences in the size ranges of strains with different ITS-types and pointed out that ITS-A was only present in 2009 samples (Table 6.5, Fig. 6.6).

Table 6.6 illustrated the genotypic diversity of strains screened both for ITS-types and microsatellite pattern. The lowest genotypic diversity that characterised POP_1 (Table 6.5) was due to high similarities between microsatellite pattern of strains exhibited ITS-A/B genotypes (Table 6.6) whereas POP_2, that presented a highest genotypic diversity (Table 6.5), showed six genotypic groups similarities. The POP_2 groups were composed of strains gathered as a function of the bloom period from which they belong but independently of their ITS-type.

Among the strains sampled in 2008, marked differences existed in the distribution of ITS-types between POP_1 (106 strains assigned) and POP_2 (29 strains assigned) (Fig. 6.7). A total of 94 strains out of the 106 in POP_1 showed an ITS A/B B-dominant type and only one of the ITS B-type. In contrast only nine out of the 29 strains in POP_2 were of the ITS A/B B-dominant type and 17 of the ITS B-type. ITS-types were available for six out of the 13 Intermediate strains (STRUCTURE); the types and distribution of these types reflect the distribution of the ITS-types over the two populations taken together (the two columns to the left in Fig. 6.7) and so

does the distribution of ITS-types over those strains for which microsatellite patterns were not obtained (the fourth column in Fig. 6.7).

Figure 6.7 showed also a marked difference between the distribution of ITS-types over the POP_2 of the 2008-sample (29 strains assigned) and of the 2009-sample (93 strains assigned). Sequences of the ITS A-type were rare among the 29 strains of POP_2 in the 2008 sample whereas the majority of the 93 strains assigned to POP_2 in 2009 possessed ITS A-type sequences (23 strains with ITS A-type, 28 with ITS A/B A-dominant type and 8 with ITS A/B AB equally dominant).

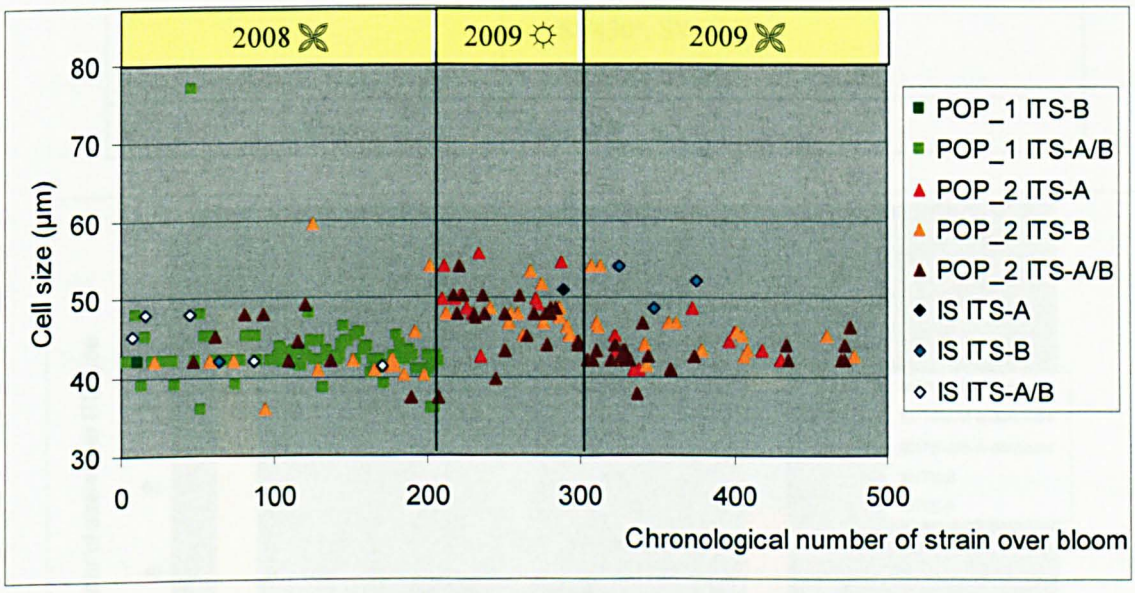


Figure 6.6: Cell size distribution of Pseudo-nitzschia multistriata *strains over 2008 and 2009. The plot is the same as in Fig. 6.5, but only those strains are shown for which ITS data were available (N= 235). Population assignment according to STRUCTURE; rectangles indicate strains assigned to POP_1; triangles, those to POP_2; diamonds, those to Intermediate strains (IS). Legend for colour coding according to ITS-types as determined in Chapter 4.*

Table 6.6: Genotypic groups of pooled strains ($N_{STRAINS}$=235 strains. See Fig. 6.6) according to microsatellite population and ITS pattern. POP_1 is represented in row two and possessed a low genetic diversity due to ITS-A/B type genotype similarities (N_{GROUP}= 11 groups). POP_2, row three, presented a gathering of six genotypic similar groups, mixed over blooms (autumn 2008 (bold red), early-summer 2009 (orange), autumn 2009 (brown)) and ITS-types (A (°), B (') and A/B (*)). Group of Intermediate strains is indicated on row four and exhibits 100% genetic diversity (N_{GROUP}= 0 groups).

Population	ITS-type	Genotypic groups
POP_1	A/B	N_{GROUP}=11
POP_2	A. B. and A/B	N_{GROUP}=6 ▪ **SY249', SY252', SY253'** ▪ **SY378°, SY328*** ▪ SY399', SY402' ▪ SY367', SY384', SY394', SY400', SY420', SY434*, SY452*, SY450', SY465', SY487', SY513', SY514'., SY578*, SY580*, SY586' ▪ SY430°, SY432° ▪ SY443°, SY467*
IS	None	N_{GROUP}=0

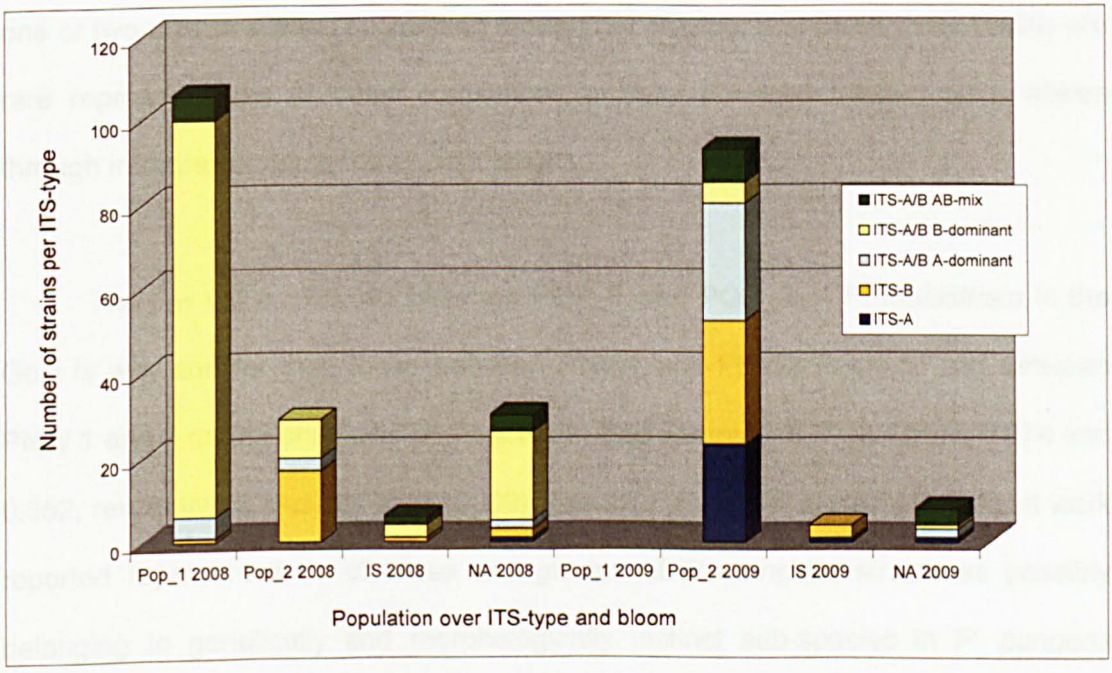


Figure 6.7. Bar chart showing distribution of ITS-types over numbers of strains in each of seven categories. Only those strains are included for which the ITS-type was determined. The height of each column is proportional to the number of strains

assigned to that category. Columns 1-4 include those strains that were sampled in 2008; columns 5-8 contain the strains collected in 2009. The 1st and 5th columns include strains assigned to POP_1 by STRUCTURE; those in the 2nd and 6th columns to POP_2, and those in 3rd and 7th columns to the Intermediate strains (IS). The 4th and 8th columns include the strains for which the ITS-type has been obtained but for which microsatellite patters were lacking or incomplete. The 5th column is empty because POP_1 was absent in 2009.

6.4. Discussion

The microsatellite data indicate that *Pseudo-nitzschia multistriata* sampled in the GoN is composed of two genetically divergent populations (POPs_1 and _2). The two populations coexisted in 2008, but in 2009, one of these (POP_1) was essentially gone. Some evidence exists for the presence of other populations. Two out of the 13 Intermediate strains were positioned well above the main swarms of strains representing the two populations (Fig. 6.3a and 6.3b) and a third strain was placed well behind the swarm in Fig. 6.3a (see Fig. 6.3b). These three strains possessed one or two unique alleles, suggesting though not proving, that these three strains are rare representatives of other populations or they obtained these unique alleles through introgression from other populations.

The F_{ST} value of 0.091 between POP_1 and POP_2 of *P. multistriata* in the GoN is way smaller than those between PNW1 and PNW2 (0.4-0.5) and between PNW 1 and 2 on the one hand and the North Sea samples of *P. pungens* (0.274 and 0.352, respectively; (Adams et al., 2009); see also (Evans et al., 2005)). Recent work reported representatives of these two groups of *P. pungens* strains as possibly belonging to genetically and morphologically distinct sub-species in *P. pungens* (Casteleyn *et al.*, 2008). Nevertheless, these divergent subspecies were still able to hybridise as demonstrated by the presence of natural hybrids and by circumstantial

evidence that these hybrids were fertile. The same accounted for the divergent geographic populations (F_{ST} >0.4) of the benthic freshwater pennate diatom *Sellaphora capitata* (Evans et al., 2009). From a population genetic perspective, one could imagine a continuum with on the one side single populations snugly in HWE and on the other extreme, different species, or multiple populations each of them in HWE, but reproductively incompatible and genetically distant from one another. In this example, the subspecies in *P. pungens* - or PNW1 and PNW2 - and the geographically isolated populations of *S. capitata* are much further on their way to becoming different species than the genotypically different groups of strains within PNW1 of *P. pungens* and the two populations in *P. multistriata*.

The results presented here and those by Evans *et al.* (2005) and Adams *et al.* (2009) indicate that hydrography and geography affect population structure within *Pseudo-nitzschia* species. The small F_{ST} values among the oceanic samples within PNW1, among the coastal and oceanic samples of PNW2 (Adams et al., 2009) and among the collections of North Sea strains (Evans et al., 2005) may be due to ongoing mixing in the eddy system off the coast of Vancouver Island or in the shallow, eutrophic epicontinental North Sea with its strong currents and eddies, driven by wind, tides and the Atlantic circulation.

The F_{ST} value between the two populations of *P. multistriata* in the GoN is comparable to that observed by Adams *et al.* (2009) between the coastal sample and the oceanic samples of *P. pungens* PNW1 (0.162-0.089). Adams *et al.* (2009) suggested that the sub-populations of *P. pungens* are potentially maintained by geographical isolation and oceanographic fronts. The Tyrrhenian Sea also offers ample opportunity for temporal geographic isolation. It constitutes a deep basin with oligotrophic surface water, fringed by narrow, shallow, nutrient-rich pockets (Astraldi

and Gasparini 1994), and *Pseudo-nitzschia multistriata* is probably restricted to these pockets. Current patterns in the Tyrrhenian Sea are much weaker than those in the North Sea (Astraldi & Gasparini, 1994). Gyres in bays and gulfs may foster the persistence of one or more populations in each of these coastal pockets, at least over the growing season, comparable to the different sub-populations of *P. pungens* found in and outside Barkley Sound, Vancouver Island by Adams *et al.* (2009). However, the co-occurrence of *P. multistriata* POP_1 and POP_2 in the same samples indicate that geographical isolation is not strictly necessary, or at least does not need to be continuous, to maintain population genetic differentiation within a planktonic diatom.

In spite of their genotypic distinctness, the two populations are not entirely isolated from one another as demonstrated by a low inbreeding coefficient (F_{IS} = -0.236; F_{IS} = -0.247 without the group of Intermediate strains), as well as attenuated gene flow (F_{ST} = 0.077; small number of migrants: Nm = 3.014). Illustrative for this gene flow are the eleven strains out of a total of 13 that could not be assigned to either population. These eleven strains were found in an intermediate position between the two swarms of data points representing the two populations (see Figs 6.3 and 6.4). Their intermediary genotypes suggest that these eleven strains are hybrids of the two populations. Adams *et al.* (2009) obtained similar results in their population genetic study of *P. pungens.* Several strains in their collection could not be assigned to either one of the two populations and the authors interpreted these strains as being inter-population hybrids.

The higher allele diversity and genotypic diversity in the sample of POP_2 (viz. those of POP_1) could be due to far larger number of strains sampled from POP_2 (227 viz. 122 from POP_1), or it may result from the different ways the

populations were sampled. Strains assigned to POP_2 have been sampled from the 2005- and 2007- cohorts and in three different periods (autumn 2008, summer 2009 and autumn 2009). Instead, strains assigned to POP_1 belonged exclusively to the 2005-cohort, and were detected only in the autumn of 2008. Genotypic diversity in the autumn sample of 2008 (93.8%) and the summer sample of 2009 (88.3%) are markedly higher than the overall genotypic diversity of POP_2 (72.7%) and the autumn samples 2009 is markedly smaller (58.5%), suggesting that there is an effect of sample period. Alternatively, the higher allele- and genotype frequencies might be due to the different histories of these populations.

A possible biological explanation for the higher genotypic diversity of POP_2 is that a prolonged period of clonal reproduction of a population in the field reduces genotypic diversity because chance-survival leads to a progressive decrease in the number of surviving clones in the population over time. The 2005-cohorts of POP_1 and POP_2 lived through three years of clonal reproduction before being sampled in 2008, whereas of the 2007-cohort of POP_2 lived through only two years of clonal reproduction before being sampled in 2009. If prolonged clonal reproduction reduces genotype diversity the POP_2 of 2009-summer with 88.31% should exhibit a higher genotypic diversity than the 2009-autumn samples of POP_2 (58.47%). However POP_2 2008 samples, expected to belong to 2005-cohort, showed a higher genotypic diversity than expected (93.8%). The most probable reason for explaining the high POP_2 2008-autumn genetic diversity is to take into account the different sampling efforts. It appears from the results that the lower the sampling effort, the higher the diversity. Indeed, in 2008-autumn the number strains sampled were 32 for a genotypic diversity of 93.8%, whereas in 2009-early summer and -autumn samples

the sampling effort were respectively of 77 strains, for a genotypic diversity of 88.3%, and 118 strain (58.5%).

Another hypothesis that may explain the high POP_2 2008-autumn genotypic diversity is an asynchrony in the sex-event of strains belonging to different ITS-types. POP_2 2008-autumn sample may be composed of strains both from 2007-cohort (of high genotypic diversity) and 2005-cohort (of low genotypic diversity); In that case the strains from 2007-cohort, of a medium cell size, should belong to sex-event induced at the very beginning of 2007 bloom; Regarding ITS-type this hypothesis also implies an asynchrony in the sexual stage between strains showing different ITS-type; In 2007, POP_2 ITS B-types strains are sexually fertile first, inducing a F1 progeny of ITS B-types, monitored among POP_2 2008 samples, then, in a second time, at the end of 2007-bloom, POP_2 ITS A-types should be fertile and be monitored in 2009-samples; This hypothesis may confirm the ITS pattern monitored in POP_2 2008 and 2009 samples, however this theory can be falsified in 2009 when no ITS-types variation were observed between 2009-early summer and -autumn samples, period during which sex event may have occurred.

A third hypothesis considers the in/out flux of populations in the Gulf of Naples. In 2008 the GoN may have received an influx current composed of cells belonging to different cohort of POP_2; In 2009, due to cell adaptation/selection (Pfrender & Lynch, 2000), part of those clonal strains may have extincted and part may have been flushed out from the GoN by out-flux currents. This phenomenon may explain the decrease of genotypic diversity observed between POP_2 2008-autumn, 2009-early summer and 2009-autumn at LTER-MC; however, because of MacDonald and Pfitzer rule (Chapter 1) the POP_2 2008-autumn should be

considered as the parental strain and therefore may not possess a such elevated clonal diversity.

A complete disappearance of an entire population, such as POP_1 from the LTER-MC in 2009, has not been reported in the other studies on *Pseudo-nitzschia* cited here. Adams *et al.* (2009) sampled their *P. pungens* populations over two years (2005-2006), but these authors did not sample the same sites, or regions, in consecutive years. Such a disappearance can be explained in two not mutually exclusive ways. First, within the GoN, the 2005-cohort of POP_2 may have been far more successful in generating offspring in 2007 than the 2005-cohort of POP_1. Second, both populations may have been flushed out of the GoN and been replaced by a similar POP_2. Since flushing out is a mechanical process it should affect both populations in the same way. Winter storms can replace the coastal eutrophic water with nutrient-poor water from the open Tyrrhenian Sea (Astraldi & Gasparini, 1994). If it is the case of a differential reproduction success, the 2005- and 2007-cohorts of POP_2 (sampled in 2008 and 2009) should be genotypically highly similar, whereas in the case of total replacement the 2005- and 2007-cohorts of POP_2 (sampled in 2008 and 2009) could be genotypically dissimilar.

The availability of ITS patterns and microsatellites provide insight in the population structure of *P. multistriata* that transcends insights that can be obtained from each of these data sets independently. The marked differences in the proportions of ITS-types between the two groups of strains as determined by STRUCTURE further support the biological reality of these two groups as belonging to two distinct populations of *P. multistriata* in the GoN. The ITS patterns also provide insight in what happened to POP_2 between 2008 and 2009.

The proportions of ITS-types within the populations in the 2008 sample are not unusual under the assumption of freely interbreeding individuals. Within the small sample of 2008-strains assigned to POP-2, the large majority of the strains exhibit pure ITS type-B or ITS type-A/B B-dominant and just a few strains exhibit ITS type-A/B A-dominant. This set of strains could be the offspring of a parental cohort in which the individuals possess many more copies of the ITS B-type than of the ITS A-type, and with many parents of the pure ITS B-type. Within the larger sample of 2008-strains assigned to POP_1, again, the vast majority exhibits ITS type-A/B B-dominant and just a few strains are of ITS type-A/B AB-mix. This set of strains could be the offspring of a parental cohort in which the individuals possess many more copies of the ITS B-type than of the ITS A-type, and most parents are of the ITS-A/B B-dominant.

Yet if the small sample of strains gathered in 2008 and assigned to POP_2 is representative of the 2005-cohort of POP_2 then that cohort cannot be the parental cohort, at least not in its entirety, of the 2007-cohort of POP_2 as sampled over the summer and autumn of 2009. The proportion of ITS-B copies has risen drastically from the 2008-sample to the 2009-sample. Moreover, a parental generation in a single population of freely interbreeding individuals would never be able to generate such a large proportion of F1 cells of the pure ITS A-type side by side with an equally large proportion of F1 cells of the pure ITS B-type.

Thus, two genetically distinct populations of *P. multistriata* occurred sympatrically in the GoN in 2008. POP_1 must have been flushed out in its entirety in the winter of 2009. POP_2 may have been flushed out as well and replaced by at least another sub-populations assignable to POP_2, but with markedly different proportions of ITS-types. Otherwise, the peculiar proportions of ITS-types in the

summer and autumn samples of 2009 go unexplained. Alternatively, a part of this POP_2 may have persisted in the GoN, but then there must have been massive immigration of cells assignable to POP_2, but with a large proportion of ITS A-types.

Evans *et al.* (2009) evaluated the population genetic structure of a freshwater diatom over six waters bodies and four countries (Belgium, Scotland, England and Australia). The strains belong to the same species: the ITS-types were similar and with absence of biogeography separation and sexual reproduction fences. The population genetic structure show however a sub-division in five populations: one related to each biogeographical region, with the exception of two for Belgium. The populations, the heterozygosity and the ITS-types should strongly be influenced by the rather non-homogeneous sample size (larger in Belgium than the other localities). Chapter 6 showed that *P. multistriata* strains in the GoN and over two years belonged to the same species because of the absence of reproductive isolation (Chapter 7 and (D'Alelio et al., 2009b)) independently of their ITS-types that presented important internal variations (Chapter 4 and (D'Alelio et al., 2009a)). The whole genotypic diversity (68.2%) found in the GoN was far lower than Evans *et al.* (2009), indicating a higher number of clonal strains in the GoN samples, with a rich genotypic diversity for pure ITS-types (ITS-A (95.8%) and ITS-B (74.5%)) rather than ITS-A/B (69.8%), providing evidence for genetic structure. However the three potential ITS-genetic populations did not correspond to the two genetic populations found in Chapter 6. Populations of the GoN presented an excess of heterozygosity which mardekly was similar to the work of Evans *et al.* (2009). As firstly hypothesised by Evans *et al.* (2009) the populations of the GoN may co-exist in the same water body with a limited dispersal of cells; however, in 2009, the hypothesis should be

invalidated by the disappearance of one population. The second hypothesis proposed by the author is the "failure of cells to persist upon arrival in a pond, perhaps because of the ecological superiority of the resident population". This second theory implies that POP_1 is the incomer and POP_2 is the resident population; the genotypic diversity of the former was by far lower than POP_2 indicating that POP_2 may perform better in the ecological water body in which they have been sampled than POP_1.

The scenarios leave us with a strange paradox: genetically different populations of *P. multistriata* exist in sympatry in an environment that shows massive disturbances, extirpating entire local populations and replacing them through immigration. How can multiple populations retain genetic distinctness in the face of such dynamic and radical exchange? One explanation might be the timing of the sexual reproduction, which may be offset from one another. In theory, even populations occurring in sympatry might retain their integrity because of, for instance, a slightly offset timing of their sexual reproduction. The sexual phase of *P. multistriata* is believed to be relatively brief in comparison to the two years of mitotic divisions to develop from initial cells to sexually mature ones (D'Alelio *et al.*, 2009a). Although sexual reproduction between strains of opposite mating types can be achieved in the lab, actual field observations of the sexual stage in *Pseudo-nitzschia* species are rare, but in those cases where sexual reproduction has been observed, it seemed to be synchronized and taking just a few days. If settings of the environmental clues to initiate sexual reproduction differ even slightly among populations, then they can occur in sympatry, but have a low probability of interbreeding. Whether the disappearance of POP_1 in the 2009 samples is due to a difference in the reproductive success between the two populations is not really

relevant for explaining the observed results because POP_2 as sampled in 2008 could not have generated POP_2 as sampled in 2009, anyway.

CHAPTER SEVEN

INHERITANCE OF INTERNAL TRANSCRIBED SPACERS AND MICROSATELLITE ALLELES DURING SEXUAL REPRODUCTION

7.1. Introduction

Diatoms feature prominently in pure and applied research projects because these photoautotrophic unicells grow extremely fast and produce various organic compounds and fine silica structures under ambient conditions (Kooistra *et al.*, 2007; Wenzl *et al.*, 2008). Many of these research projects would benefit greatly from not only growing their diatom strains of interest, but from controlling their life cycle as well. First, there is an absolute need for this because average cell size decreases with ongoing mitotic divisions as a result of an architectural constraint in their compound silica cell wall, thus restricting the lifetime of research strains to just a few years. Sexual reproduction and auxospore formation are needed to restore the initial cell size, thereby keeping at least the lineages indefinitely. Moreover, control of the sexual cycle in diatoms allows breeding programmes and selection for desired combinations of traits in the offspring. Crucial to such breeding programmes is an understanding of how meiosis proceeds and how traits are inherited. Microsatellite loci and their various alleles constitute ideal markers for doing just that.

Pseudo-nitzschia multistriata the species in the focus of my research is a raphid pennate planktonic diatom. It exhibits a diplont, biennial (2-years) life cycle, with sexual reproduction occurring at the end of the second year (D'Alelio et al., 2010). Sexual reproduction is dioecious, with cells of the opposite mating type

aligning in pairs and becoming gametangia, thus conducting meiosis together. During the initial stage of meiosis the parental chromosomes duplicate (2N to 4N), team up in homologous pairs and engage in crossing-over (Griffiths *et al.*, 1993). At the anaphase of the first meiotic division, each pair of homologous chromosomes segregates over two diploid nuclei. Two possibilities: (i) One of these nuclei is believed to go pycnotic and disappear (Round et al., 1990). During the second meiotic division, the two chromatids of the chromosomes in the remaining nucleus split at their centromeres, and the resulting haploid sister chromosomes segregate during anaphase over two haploid gametes per gametangium; or (ii) After the first meiotic division both diploid nuclei survive and after the second meiotic division only one out of two haploid nuclei survives (Mann *et al.*, 2003; Round *et al.*, 1990). The two gametes produced by one gametangium (designated the paternal or "-" gametangium) move in an amoeboid fashion to the two gametes produced by the other gametangium (designated the maternal or "-" gametangium) and fuse with them to form two zygotes that remain attached to the parental cell wall. Each zygote inflates anisometrically to form an auxospore from which the initial cell emerges (Kooistra et al., 2007).

In previous chapters in this thesis, molecular markers were developed to assess population genetic structure of *P. multistriata*. Therefore, a total of 467 strains were collected at the long-term ecological research station MareChiara (LTER-MC) in the Gulf of Naples (GoN) in 2008 and 2009. The molecular data included the internal transcribed spacers (ITS-1 and ITS-2) of the nuclear rDNA cistrons from 265 of the 467 strains and seven microsatellite fingerprints from 412 of the 467 strains. Two genetically markedly distinct populations were identified among the microsatellite

fingerprints. These populations also showed distinct proportions of the various ITS-types recognized. Distinct proportions in the ITS-types were also observed within populations from one year to the next.

The results generated several new questions. Does sex occur between strains belonging to different ITS-types, and if so, then how do ITS-types inherit? Do microsatellites inherit according to Mendelian rules? Do these markers exhibit null alleles, and if so, can they be uncovered by genotyping the F1? Is sexual reproduction successful between strains belonging to different populations as identified with microsatellites?

These issues can be addressed by screening microsatellite fingerprints and ITS-types of sufficient numbers of F1 cells derived from parents of known genotypes. Initial F1 cells are recognized easily from their parents because they are taller [80 to 70 μm] than their parental cells [55 to 30 μm]. Ideally, pairs of initial cells are isolated the moment they emerge from the sister pairs of auxospores, as these pairs can provide additional information about the segregation mode during meiosis and may reveal recombination events. Isolation of the auxospores themselves is unfeasible, as these are flimsy structures.

A set of five parental strains (two females and three males) was chosen to cover the two populations as inferred from the microsatellite markers (Chapter 6), different ITS-types (Chapter 4) and different sampling periods (Appendix 1). Crosses were performed between parental strains of the ITS A-type and B-type and between the ITS A-type and ITS A/B-type to compare observed proportions of ITS-types in the offspring with expected proportions. In addition, the parental strains were selected such that their microsatellite genotypes allow assessment of:

(i) Comparison of observed with expected genotype diversity among the F1,

(ii) Compliance of the F1 genotype distribution with Mendelian inheritance laws,

(iii) The parental contribution to the F1 genotypes,

(iv) The contribution of the progeny to the genotype diversity of a population,

(v) The existence of null alleles.

7.2. Materials and Methods

The five strains: SY017, SY138, SY278, SY378 and SY379 used for the crossing experiments were isolated from the early summer and autumn of 2008 and 2009 (Table 7.1). Cell sizes ranged between *ca.* 46.8 and 25 μm during the crossing experiments, which was within the cell size window for sexual reproduction (D'Alelio et al., 2010). Mating-types were assigned arbitrarily "–" for strains producing passive female gametes and "+" for those producing active male gametes that fused with female gametes, based on the results of pilot experiments in which strains of known genotypes and ITS-types were crossed in all possible pair-wise combinations.

Parental strains were grown in F/2 filtered medium (Sigma Aldrich S.r.l. at 36 psu final salinity concentration) at 15 °C under a photoperiod of 12:12h L:D with sinusoidal light regime at a maximum irradiance of 80 μmol photon$\cdot m^{-2} \cdot s^{-1}$ provided by cool-white fluorescent tubes (Philips T94 Eindhoven, the Netherlands, see Chapter 3).

Table 7.1: Parental strains, used in the crosses; 'Sample period' denotes year and season (early-summer (☼) or autumnal (✂) in which strains were obtained, 'MT' the, mating type (Female (-) and Male (+)), 'Crosses' the crossing in which strains were involved (consecutive numbers indicate the different crosses), ITS-types (Chapter 4) and microsatellite population (POP_1, POP_2, Chapter 6).

Strain	Bloom	MT	Crosses	ITS-type	Microsatellite Population
SY017	2008 ☼	-	1,2	A	1
SY138	2008 ✂	+	2,3,4	B	2
SY278	2008 ✂	+	1	B	2
SY378	2009 ☼	-	3	A/B	2
SY379	2009 ☼	-	4	A	2

A minimum of sixty crosses has been performed for each parental couple (Table 7.2). Crosses were performed bringing together *ca.* 5000 cells·ml^{-1} of each exponentially growing parental culture in each well of a 6-well plate (Corning® Costar®, NY, USA) filled with 4 ml of filtrated F/2 medium. Cell concentration was assessed with a 1-ml Sedgewick-Rafter, counted with an Axiophot microscope (x125 magnification, Zeiss, Göttingen, Germany).

The cultures of parental couples were incubated at 18 °C, 12:12h L:D daylength, on/off regime, at 80-90 μmol photon·m^{-2}·s^{-1} irradiance provided by cool-white fluorescent tubes (Phillips TLD 36W/950). Five to eight days after the start of a crossing experiment, cells usually had produced auxospores and initial cells. Two to four tall cells deriving from one or two divisions of the initial cells were isolated randomly from each culture well under a light microscope (100x magnification; Leica Microsystems DMIL, Wetzlar, Germany) using drawn-out glass Pasteur pipettes. Each isolated cell was transferred to a well in a 12 well culture plate (Corning® Costar® CLS3513, NY, USA) each containing 2 ml F/2 filtrated medium and

incubated under the same conditions as described for the crossing experiments. A total of 251 cultures from post-initial cells and two cultures (CRD1 and CRD2) of sister-cells were obtained (Table 7.2, Appendix 3). In case of contamination by parental cells, strains were re-isolated to get a monoclonal F1 culture. After one week in the culture plates, the F1 strains were transferred in 30 ml F/2 medium (Corning® Flask, code 430639, Corning Inc., NY, USA) and incubated for one more week. DNA was extracted following the modified CTAB protocol described in Chapter 2. ITS-sequences and seven polymorphic microsatellites markers (*PNm2, PNm5, PNm7, PNm3, PNm6, PNm16* and *PNm1*) were analysed (see Chapter 2 for methodological details and Chapters 4 and 6 for sequences analyses, Appendix 3).

BioEdit (version 7.0.9.0, (Hall *et al.*, 1992)) was used to identify ITS-types. CEQ™2000XL (version 4.3.9, Beckman Coulter™) was used to infer microsatellite length. Microsatellites lengths were classed in length-windows determined during the stabilisation of protocol (Chapter 5). Gimlet (version 1.3.3, (Valière, 2002)) was deployed to assess genotypic diversity, GenAlEx (version 6, (Peakall & Smouse, 2006)) to estimate population genetic parameters, Genetix (version 4.05.2) to calculate linkage dis-equilibrium. To visualise genetic relationships of F1 strains and the influence of parental genotypes on the F1 generation, pair-wise distances among the microsatellite genotypes of the parents and progeny of all crosses were plotted according to results of a Factorial Correspondence Analysis (FCA) using Genetix., Genetic structure among parental and F1 genotypes was assessed using Structure (version 2.1, (Pritchard et al., 2000)). Normally, Structure is used to delineate within a sample those groups of individuals that are in HWE among one another, and that differ genotypically from other such groups in the sample. For that reason the

program separates also the offspring from one cross from offspring belonging to another cross if one or both of the parents are genotypically markedly distinct.

Statistical approaches were performed using two statistical tests: (i) the Chi^2 test (Pearson Chi-square test; Appendix 13) and (ii) the comparison of proportion test (also called Z-test; Appendices 11 and 12).

The Chi^2 test: was applied when the expected frequencies per microsatellite locus were more than three allelic combinations and estimated to (¼, ¼, and ½ for a degree of liberty of 2) or (¼, ¼, ¼, and ¼ for a degree of liberty of 3). The observed distributions of allele frequency (Chi^2_{obs}) were compared with the expected one (Chi^2_{th}) and significance of proportion differences were assessed following the test Chi^2-law of conformity. The null hypothesis (Ho) assumes that the observed distribution follows the Mendelian inheritance law; whereas the alternative hypothesis (H1) states that the two proportions are different and the inheritance pattern does not conform to Mendelian law.

- *The parameter set-up:* under bilateral distribution and an error of 5% (α = 5 %, Chi^2_{th} = 5.99, for a ddl = 2 and Chi^2_{th} = 7.81, for a ddl = 3).

- *The validation of the test*: If Chi^2_{obs} is lower than Chi^2_{th}, Ho is accepted under α. If Chi^2_{obs} is higher than or equal to Chi^2_{th}, H1 is accepted under α.

Test of comparison of proportion (Z-test): follows the Reduced Normal Table. This test was applied (i) to compare two proportions (p_1 and p_2) or (ii) when the expected microsatellite proportion at a locus was ½ and ½. The observed distributions of allele

frequency (Zc) were compared with the expected one ($U_{\alpha/2}$). Common proportion (p_c) was calculated as the sum of a particular genotype (n = N x p) over the sum of strains analysed (N):

➢ $$p_c = \frac{N_1 p_1 + N_2 p_2}{N_1 + N_2}$$

The observed frequency was calculated as follow:

➢ $$Zc = \left| \frac{(p_1 - p_2)}{\sqrt{\frac{(p_c (1 - p_c))}{p_1} + \frac{(p_c (1 - p_c))}{p_2}}} \right|$$

The null hypothesis (Ho) assumes that (i) the two proportions tested are equal or (ii) that the observed distribution follows the Mendelian inheritance law; whereas the alternative hypothesis (H1) settle that (i) the two proportions tested are different or (ii) the two proportions are different and the inheritance law is not conformed to Mendelian's law.

- *The parameters set-up:* under bilateral distribution and an error of 5 % (α = 5 %, $U_{\alpha/2}$ = 1.96) or 1 % (α = 1%, $U_{\alpha/2}$= 2.57).

- *The validation of the test*: number of strains tested (n > 30) and its relative proportion (np > 5). If Zc value is higher than $U_{\alpha/2}$, Ho can not statistically be accepted under alpha and H1 is validated.

7.3. Results

All four combinations of parental strains generated auxospores and F1 initial cells, and all of the latter divided clonally as demonstrated by chain formation. Visual examination showed abundant auxosporulation within two days in Cross_1 and Cross_2 and far less auxosporulation in Cross_3 and Cross_4. The vigorous crosses were between a parent from POP_1 and one from POP_2, but both involved strain SY017 and were between partners of smallish cell size. The less vigorous crosses were between parents from within POP_2, but also involved at least one relatively tall partner strain. In any case, reproduction success did not depend on the ITS-type (Table 7.1).

Results of isolation of the F1 cells revealed a low survival frequency (Table 7.2). Survival of the F1 was not linked to whether the parents belonged to the same or different populations, or whether they were isolated in the same or in different years. However, Cross_3, the one with the lowest survival (8.25%) of the isolated F1 cells involved a parental strain of the ITS A/B-type (SY378).

Internal Transcribed Spacer (ITS): Of the 253 strains established from the offspring of the crossing experiments, 122 strains were successfully analysed for their ITS-types pattern (Table 7.3). ITS-sequences analysis and reading efficiencies depended on the crossing conducted, as several of the ITS electopherograms were unreadable or the sequencing of the PCR products was unsuccessful. Progeny from the Cross_2 and Cross_4, involving ITS-A x ITS-B parent types, were 100% ITS A/B-type, as expected. However Cross_1, involving ITS-A x ITS-B parent types, resulted in three out of 34 F1 strains of the ITS A-type. Cross_3 (ITS-A/B x ITS-B) resulted in nine F1 strains of the ITS A/B-type and two of the ITS B-type (Table 7.3).

*Table 7.2: Results of crossing experiments. 'Cross' denotes sequential number of the cross; 'Strain (-)' and 'Strain (+)' denote parental strains involved in the cross and their mating type; 'Size (µm)' denotes the average cell length measured over the apical axis; '# Crosses performed' indicates the number of wells in Petri plates in which strains were brought together in pairs; 'Isolated' signifies the total number of F1 cells that were isolated from those wells, and 'surviving' means those isolated F1 cells that survived and grew into monoclonal strains. ** indicate the estimated size value ($S_f = S_i - n_{months} \cdot C_{regression}$) see Appendix 9.*

Cross	Strain (-) (+)	Size (µm)	# Crosses performed	# Strains surviving / isolated (survival rate)	Strain Codes	Date of experiments and isolation
1	SY017 SY278	38-34 36-32	150	62 / 218 (28.4%)	CRA-F CR1-30B CS1-120B	July 2009 Sept. 2009 Oct. 2009
2	SY017 SY138	34 25	60	95 / 307 (30.9%)	CS121-180D	Oct. 2009
3	SY378 SY138	45** 25	60	26 / 315 (8.25%)	CS181-240D	Nov. 2009
4	SY379 SY138	46.8** 25	60	72 / 307 (23.5%)	CS241-300D	Nov. 2009

Table 7.3: Numbers of ITS-types (A, B, and A/B) among the offspring of the four crosses conducted. See Tables 7.1 and 7.2 for information about parental strains.

	Cross 1	Cross 2	Cross 3	Cross 4
ITS-types of parental strains	A x B	A x B	A/B x B	A x B
Number of F1 strains exhibiting readable sequences	34	31	11	46
A	3 ; 8.8 %	0 ; 0 %	0 ; 0 %	0 ; 0 %
B	0 ; 0.0 %	0 ; 0 %	2 ; 18.2 %	0 ; 0 %
A/B	31 ; 91.2 %	31 ; 100 %	9 ; 81.8 %	46 ; 100 %
Number of F1 strains exhibiting unreadable sequences	28	64	15	26
Sum of F1 strains examined	62	95	26	72

Microsatellite markers: Of the 253 strains established from the offspring of the crossing experiments, 246 strains were successfully analysed for microsatellite markers. The analysis and reading efficiencies differed marginally among the crosses (Cross_1 (100%), Cross_2 (96.8%), Cross_3 (100%) and Cross_4 (91.7%)).

Microsatellite genotype diversity: The expected number of possible F1 genotypes (N_G) was highest in Cross_1 because both parents were heterozygous for most of the loci (Table 7.4). Yet, the actual number of observed genotypes in the F1 was, of course, much lower than N_G because especially in the case of Cross_1 the 62 strains isolated from that cross did not permit detecting all possible 864 genotypes (see Tables 7.5 and 7.6). All F1 strains, except one (CS272A2), exhibited genotypes as expected within the spectrum of possible genotypes of the cross, given the genotypes of the parents. The observed genotypic diversity was lowest among the sampled F1 strains of Cross_4 (43.48%). In contrast, the samples of the F1 offspring of Crosses_1,_2 and_3 exhibited genotype diversities of 85.49%, 87.10% and 90.91%, respectively (Table 7.6). As expected, the observed genotypic diversity of the F1 was independent from the parental ITS-types (Table 7.6, Gimlet Appendix 10).

Strain CS272A2 differed from all other F1 strains of Cross_4 in that it exhibited the same microsatellite genotype as its parental strain SY379 (Table 7.7). The strain was an F1 because its average length over the apical axis was 74.6 (± 1.06) μm, while that of the tallest parental strain, SY379, was markedly lower (46.8 μm). Moreover, the ITS of strain CS272A2 was of the A/B-type, as expected for the offspring of an ITS A-type and an ITS B-type parental strain.

Table 7.4: Microsatellite alleles at each of the seven loci of the parental strains utilised in the four crosses; cross sequential number (C), the mating type (T) and the code of parental strains (Strains), the microsatellite population to which they belong (P; Chapter 6), their microsatellite pattern for the seven loci amplified. A White box highlight the alleles that + and – parents share; blue the specific alleles of the male (+) genotype and pink those of the female (-) genotype.

C	T	Strains	P	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
1	-	SY017	1	A	175	175	238	242	262	262	205	208	261	267	325	337	127	127
	+	SY278	2	B	175	187	238	242	258	262	205	208	267	271	325	337	115	127
2	-	SY017	1	A	175	175	238	242	262	262	205	208	261	267	325	337	127	127
	+	SY138	2	B	175	187	238	238	258	258	208	208	255	255	307	337	117	127
3	-	SY378	2	A/B	175	187	238	238	262	262	208	208	249	249	325	337	115	117
	+	SY138	2	B	175	187	238	238	258	258	208	208	255	255	307	337	117	127
4	-	SY379	2	A	175	175	230	238	262	262	208	208	267	267	319	325	117	127
	+	SY138	2	B	175	187	238	238	258	258	208	208	255	255	307	337	117	127

Table 7.5: Number of all possible microsatellite genotypes of the F1 offspring for the four crosses, given the parental genotypes. The numbers in the columns under a locus indicate the number of genotypes for that locus in the progeny of the particular cross. The total possible number of F1 genotypes N_G is the product of the number of allelic combinations at each locus over all loci tested, given the alleles on those loci in the parental strains (Table 7.4).

Cross	*PNm2*	*PNm5*	*PNm7*	*PNm3*	*PNm6*	*PNm16*	*PNm1*	N_G
1	2	3	2	3	4	3	2	864
2	2	2	1	2	2	4	2	128
3	3	1	1	1	1	4	4	48
4	2	2	1	1	1	4	3	48

For example if A, a, α denote alleles, then a cross AA x AA generates an F1 of only one genotypes, a cross Aa x AA generates two genotypes (Aa and AA), a cross Aa x Aa generates three (AA, Aa, aa), a cross Aa x a α four (Aa, aa, A α, a α) and so does a cross Aa x A α (AA / Aa / A α / a α).

In Cross_3 the genotypic diversity, taken into account ITS-types, was very high (100%; Table 7.6). However this genotypes diversity decreased when the ITS-types were pooled together (90.96%; Table 7.6).

Mendelian inheritance: Observed genotype frequencies of the F1 conformed to the expected frequencies at almost all loci in almost all of the F1 strains (see results of Chi^2 tests and comparison of frequencies-tests, Tables 7.8-7.11 and Appendices 11-13). Expected frequencies of the F1 were inferred from parental genotypes following Mendelian inheritance laws. Nonetheless, seven loci over the 28 analysed (7 loci x 4 crosses) showed inheritance patterns deviating from expectations. Rare, but non-significant, deviations from Mendelian law were observed on: *PNm5* at Cross_2, *PNm7* at Cross_4, *PNm6* at Cross_4, *PNm16* at Cross_1 and Cross_4 and *PNm1* at Cross_2 and Cross_3 (Tables 7.8-7.11). The rare deviations from Mendelian law could be due to the four following processes. First, inheritance of both alleles from the same parental strain at one locus occurred at 1.5% of the F1 genotypes of Cross_4 (Table 7.11). One F1 strain inherited both alleles from their maternal strain SY379 on loci *PNm7, PNm6* and *PNm16* and one F1 strain inherited both alleles from its paternal strain SY138 on locus *PNm16*. Second, allele inheritance disequilibrium was observed in Cross_1 and Cross_4. In Cross_1 (*PNm1)* the observed genotype (100% AA - 0% Aa) deviated from expectation (50% AA - 50% Aa) and in Cross_4 (*PNm7* and *PNm6*) the observed genotypes (98% Aa - 2% AA) deviated from the expected (100% Aa; Tables 7.5 and 7.8). Third, novel alleles were rare (3.26 to 5.43% of strains at the given locus) and detected at Cross_1 (*PNm16*), Cross_2 (*PNm5* and *PNm1*) and Cross_3 (*PNm1*) (Tables 7.8-7.10).

*Table 7.6: F1 Microsatellite genotypic diversity in the different crosses and in function of their ITS-type. The number of F1 strains tested (N), the number of unique genotype among N (G/N) and its proportion, the genetic diversity per ITS-type (G/N (ITS-type)) and its proportion. The * indicates F1 particular genotype linkage (Appendix 10), with green, belonging to ITS-A/B type and orange to ITS-B type.*

	N	G/N	G/N (ITS-A)	G/N (ITS-B)	G/N (ITS-A/B)
Cross_1 ITS- A x B	62	29/34 ; 85.29%	3/3 ; 100%	None	26/31; 83.87%
Cross_2 ITS- A x B	92	27/31 ; 87.10%	None	None	27/31; 87.10%
Cross_3 ITS- A/B x B	26	10/11 ; 90.91% *CS233C & CS225C	None	2/2 = 100%	9/9 ; 100%
Cross_4 ITS- A x B	66	20/46 ; 43.48%	None	None	20/46 ; 43.48% * SY379 & CS272A2
Whole	246	85/121; 70.25%	3/3 ; 100%	2/2 = 100%	81/117 ; 69.23% One match

Table 7.7: ITS-type and microsatellite alleles on the seven loci of parental strains SY379 and SY138 and of one of their offspring strain (CS272A2, Cross_4). For mean of the coloured boxes, see Table 7.4.

	Strains	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
P	SY379	A	175	175	230	238	262	262	208	208	267	267	319	325	117	127
P	SY138	B	175	187	238	238	258	258	208	208	255	255	307	337	117	127
F1	CS272A2	A/B	175	175	230	238	262	262	208	208	267	267	319	325	117	127

Table 7.8: Mendelian inheritance estimation in Cross_1. The locus tested (locus), the parental genotype characteristics (strain code, allele length on site 1 and 2 at a given locus (A1, A2)), the progeny genotype (allele length on A1 and A2), the specific and total number of F1 genotype combination found (# F1 genotypes), the statistical approaches (frequencies comparison test and Chi² test (ddl=3 and 4)) under alpha 5%. Ho signifies null hypothesis accepted; H1, null hypothesis rejected; Δ informs of rare event. Blue and pink boxes indicate male and female-specific alleles, respectively; green indicates double peaks occurrence at a given allele. Statistical tests, see Appendices 11 and 13.

CROSS_1	P genotype			F1 genotype		# F1 genotypes		statistical values		
locus	strain	A1	A2	A1	A2	specific	total	f(obs)	f(exp)	test
PNm2	SY017	175	175	175	175	26	62	0.42	0.50	Ho
	SY278	175	187	187	175	36	62	0.58	0.50	Ho
PNm5	SY017	238	242	238	238	19	60	0.32	0.25	Ho
	SY278	238	242	242	242	11	60	0.18	0.25	Ho
				238	242	30	60	0.50	0.50	Ho
PNm7	SY017	262	262	262	262	24	61	0.39	0.50	Ho
	SY278	258	262	258	262	37	61	0.61	0.50	Ho
PNm3	SY017	205	208	205	205	14	62	0.23	0.25	Ho
	SY278	205	208	208	208	14	62	0.23	0.25	Ho
				205	208	34	62	0.55	0.50	Ho
PNm6	SY017	261	267	261	267	15	61	0.25	0.25	Ho
	SY278	267	271	261	271	12	61	0.20	0.25	Ho
				267	267	19	61	0.31	0.25	Ho
				267	271	15	61	0.25	0.25	Ho
PNm16	SY017	325	337	325	325	10	61	0.16	0.25	Ho
	SY278	325	337	337	337	18	61	0.30	0.25	Ho
				325	337	31	61	0.51	0.50	Ho
				319/325	337	2	61	0.03	0.00	Δ
PNm1	SY017	127	127	127	127	62	62	1.00	0.50	H1
	SY278	115	127	115	127	0	62	0.00	0.50	H1

Table 7.9: Mendelian inheritance estimation in Cross_2. Legend (see Table 7.8). Statistical tests, see Appendices 11 and 13.

CROSS_2	P genotype			F1 genotype		# F1 genotypes		statistical values		
locus	strain	A1	A2	A1	A2	specific	total	f(obs)	f(exp)	test
PNm2	SY017	175	175	175	175	49	92	0.53	0.50	Ho
	SY138	175	187	175	187	43	92	0.47	0.50	Ho
PNm5	SY017	238	242	238	242	37	89	0.42	0.50	Ho
	SY138	238	238	238	238	49	89	0.55	0.50	Ho
				238	248	1	89	0.01	0.00	Δ
				238	240	1	89	0.01	0.00	Δ
				240	240	1	89	0.01	0.00	Δ
PNm7	SY017	262	262	258	262	87	87	1.00	1.00	Ho
	SY138	258	258							
PNm3	SY017	205	208	205	208	43	92	0.47	0.50	Ho
	SY138	208	208	208	208	49	92	0.53	0.50	Ho
PNm6	SY017	261	267	255	261	37	86	0.43	0.50	Ho
	SY138	255	255	255	267	49	86	0.57	0.50	Ho
PNm16	SY017	325	337	307	325	24	90	0.27	0.25	Ho
	SY138	307	337	307	337	22	90	0.24	0.25	Ho
				325	337	24	90	0.27	0.25	Ho
				337	337	20	90	0.22	0.25	Ho
PNm1	SY017	127	127	117	127	43	92	0.47	0.50	Ho
	SY138	117	127	127	127	45	92	0.49	0.50	Ho
				119	127	3	92	0.03	0.00	Δ
				117	117	1	92	0.01	0.00	Δ

Table 7.10: Mendelian inheritance estimation in Cross_3. Legend (see Table 7.8). Statistical tests, see Appendices 12 and 13.

CROSS_3	P genotype			F1 genotype		# F1 genotypes		statistical values		
Locus	strain	A1	A2	A1	A2	specific	total	f(obs)	f(exp)	Test
PNm2	SY378	175	187	175	175	5	26	0.19	0.25	Ho
	SY138	175	187	175	187	18	26	0.69	0.50	Ho
				187	187	3	26	0.12	0.25	Ho
PNm5	SY378	238	238	238	238	26	26	1.00	1.00	Ho
	SY138	238	238							
PNm7	SY378	262	262	258	262	26	26	1.00	1.00	Ho
	SY138	258	258							
PNm3	SY378	208	208	208	208	26	26	1.00	1.00	Ho
	SY138	208	208							
PNm6	SY378	249	249	249	255	26	26	1.00	1.00	Ho
	SY138	255	255							
PNm16	SY378	307	337	307	325	5	25	0.20	0.25	Ho
	SY138	325	337	307	337	4	25	0.16	0.25	Ho
				325	337	12	25	0.48	0.25	Ho
				337	337	4	25	0.16	0.25	Ho
PNm1	SY378	115	117	115	117	1	26	0.04	0.25	Ho
	SY138	117	127	115	127	6	26	0.23	0.25	Ho
				117	127	10	26	0.38	0.25	Ho
				117	117	4	26	0.15	0.25	Ho
				119	119	4	26	0.15	0.00	Δ
				117	119	1	26	0.04	0.00	Δ

Table 7.11: Mendelian inheritance estimation in Cross_4. Legend (see Table 7.8). Statistical tests, see Appendices 12 and 13.

CROSS_4	Pgenotype			F1 genotype		# F1 genotypes		statistical values		
Locus	strain	A1	A2	A1	A2	specific	total	f(obs)	f(exp)	test
PNm2	SY379	175	175	175	175	25	67	0.37	0.50	Ho
	SY138	175	187	175	187	42	67	0.63	0.50	Ho
PNm5	SY379	230	238	230	238	37	67	0.55	0.50	Ho
	SY138	238	238	238	238	30	67	0.45	0.50	Ho
PNm7	SY379	262	262	258	262	65	66	0.98	1.00	Ho
	SY138	258	258	262	262	1	66	0.02	0.00	Δ
PNm3	SY379	208	208	208	208	67	67	1.00	1.00	Ho
	SY138	208	208							
PNm6	SY379	267	267	255	267	65	66	0.98	1.00	Ho
	SY138	255	255	267	267	1	66	0.02	0.00	Δ
PNm16	SY379	319	325	307	319	16	67	0.24	0.25	H1
	SY138	307	337	307	325	10	67	0.15	0.25	H1
				319	337	13	67	0.19	0.25	H1
				325	337	26	67	0.39	0.25	H1
				307/337	319	1	67	0.01	0.00	Δ
				319	325	1	67	0.01	0.00	Δ
PNm1	SY379	117	127	117	117	21	66	0.32	0.25	Ho
	SY138	117	127	117	127	33	66	0.50	0.50	Ho
				127	127	13	66	0.20	0.25	Ho

And fourth, null alleles occur in Cross_2 *PNm1* because the F1 showed inheritance of an allele only from the paternal strain (SY138). However this phenomenon was not detected at this locus in Cross_3 and Cross_4 with the same paternal strain.

Parental contribution: The contribution of the "-" and "+" mates to the F1 genotypes was assessed for crosses in which all parental alleles of a given locus were different (Table 7.12), thus permitting unambiguous determination of parental contribution to the genotypes of the offspring. The loci tested were *PNm7* (Crosses_2, _3 and _4), *PNm6* (Crosses_2, _3 and _4) and *PNm16* (Cross_4). At loci *PNm7* and *PNm6* (Crosses_2 and _3) the female and the male contributed equally to F1 genotype. However on Cross_4 (*PNm7*, *PNm6* and *PNm16*) a minor dis-equilibrium in the Mendelian expectation was observed. The deviation of the contribution applied especially to the male for which the contribution in F1 genotype is 49.3% instead of the 50% expected from Mendelian law (Table 7.12).

*Table 7.12: Parental contribution to F1 genotypes. The three loci tested are PNm7, PNm6 and PNm16. Numbers indicate the contribution of each parental genotype to the F1 genotype and the numbers in brackets indicate the parental contribution to the F1 gamete genotype. The 'Reason' row postulates why the contribution of parental strain in Cross_4 diverges from the expected parental distribution. A * denotes that the contribution was not assessable.*

	Parental strain	*PNm7*	*PNm6*	*PNm16*
Cross_2	Male SY138	100%, (50%)	100%, (50%)	*
	Female SY017	100%, (50%)	100%, (50%)	*
Cross_3	Male SY138	100%, (50%)	100%, (50%)	*
	Female SY378	100%, (50%)	100%, (50%)	*
Cross_4	Male SY138	98.5%, (49.3%)	98.5%, (49.3%)	98.5%, (49.3%)
	Female SY379	100%, (50%)	100%, (50%)	100%, (50%)
	Reason	Non disjunction	Non disjunction	Non disjunction

Table 7.13: Genotypes of sister strains CRD1 and CRD2 raised from two initial cells produced on the same gametangium. The genotypes of parental (P) and progeny (F1) in Cross_1, the ITS-types, and the microsatellite pattern over the seven loci tested each one composed of two alleles. White box highlight the alleles that the genotypes of the parental strains have in common; blue the specific alleles of the male (+) genotype and pink those of the female (–) genotype in parental strains and the inheritance origin in F1 strain.

	Strains	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
P	SY017	A	175	175	238	242	262	262	205	208	261	267	325	337	127	127
P	SY278	B	175	187	238	242	258	262	205	208	267	271	325	337	115	127
F1	CRD1	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127
F1	CRD2	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127

Table 7.14: Auto-auxosporulation hypothesis. The genotypes of parental (P) and progeny (F1, sister-cells) strains in Cross_1, the strains code (Strains), the ITS-types, and the microsatellite pattern over the seven loci tested each one composed of two alleles. Box colours, see Table 7.13.

	Strains	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
P	SY278	B	175	187	238	242	258	262	205	208	267	271	325	337	115	127
P	SY278	B	175	187	238	242	258	262	205	208	267	271	325	337	115	127
F1	CRD1	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127
F1	CRD2	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127

Among the 62 F1 strains of Cross_1, one pair of strains, namely CRD1 and CRD2, was generated from sister-cells produced on the same gametangium. CRD1 and CRD2 were genetically identical (Table 7.13, Appendix 10). They possessed an ITS A/B-type pattern as expected for F1 progeny of ITS A-type and B-type parents (Table 7.13). In addition, they possessed identical microsatellite genotypes, which differed from genotypes of the parents (Table 7.13). The sister cells did not inherit only the genotype of SY017 because that would leave the alleles on *PNm7* and *PNm6* unexplained. Alternatively, homothallic (*i.e.* intra-clonal) auxosporulation of strain SY278 could explain the observed result (Table 7.14). According to the microsatellite patterns, homothallic auxosporulation is possible because F1 progeny could have inherited from the SY278 allele 175 from both parents on *PNm2* and allele 127 from both parents on *PNm1*. What makes this possibility unlikely is that homothallic auxosporulation could not explain the ITS A/B-type pattern of F1 progeny from parents SY278 only.

Contribution to bloom dynamics: Factorial Correspondence Analyses of pair-wise distances among parental and offspring-genotypes have been carried out on the complete set of strains (N=251, *ca.* 20 of which presented up to four missing values) and on a subset of 231 strains for which all genotypes were known. Results of the latter analysis show that F1 genotypes belonging to the same cross were spatially correlated (Fig. 7.1). The offspring of crosses sharing one of their parents were formed nearby or even partially overlapping clusters: the offspring of Crosses_2, _3 and _4 (grey, green and pink clusters, respectively) shared the same paternal strain (SY138) and the offspring of Crosses_1 and _2 (blue and grey clusters) shared the

same maternal strain (SY017). The offspring of Cross_2, (grey cluster in centre of plot), shared a parent (SY017 or SY138) with all of the other crosses. F1 strains generally grouped near or around their parental strains.

Results of STRUCTURE (K=1) analysis of the offspring-genotypes (Fig. 7.2A) revealed a single cluster subdivided into three groups corresponding to the progeny of the three crosses in which at least one of the parents was genotypically distinct (blue, Cross_1; red Cross_3; green Cross_4) and a zone of mixed genotypes corresponding to the progeny of Cross_2. Results of AMOVA showed 38% of variance among populations, thus confirming the results of the STRUCTURE analysis.

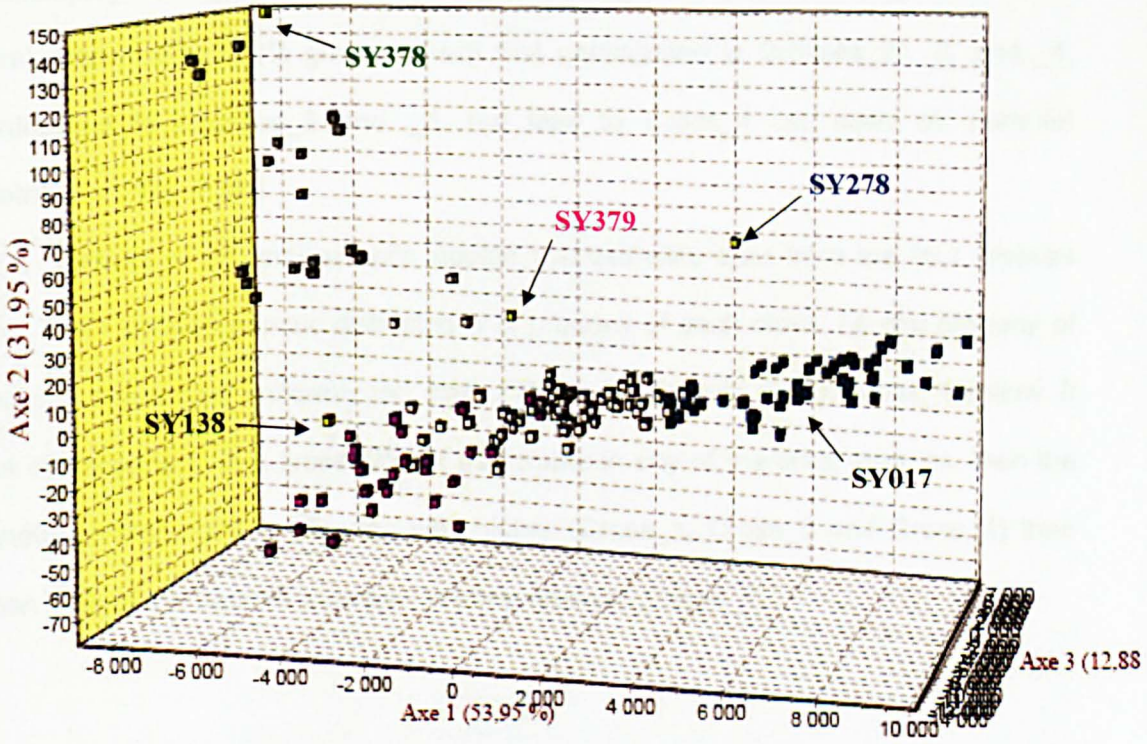


Figure 7.1: Factorial Correspondence Analysis of microsatellite genotype distances (N=251) among all strains (F1 and parents; blue cloud: Cross_1 SY017 x SY278, grey cloud: Cross_2 SY017 x SY138, green cloud: Cross_3 SY378 x SY138 and pink cloud: Cross_4 SY379 x SY138) using GENETIX. Arrows indicate parental strains, strain codes in black denote parents involved in multiple crosses, coloured strains

codes indicate parents involved in only one cross; colour corresponding to that of cloud.

STRUCTURE resolved the offspring of Cross_2 as a 'population of hybrids' with input from the other three 'groups of step-sisters' (Appendices 14 and 15); one parent SY017 also participated in Cross_1 and the other parent SY138 in Crosses_3 and _4.

Specific parental contribution was indicated on Fig. 7.2B. SY278 (blue), SY378 (½ green ½ red) and SY379 (green) involved in Crosses_1, _3 and _4, respectively, and contributed fully to the genome of the progeny. Parent SY017 (blue) participated in Crosses_1 and _2. It contributed fully to Cross_1, but less to Cross_2 (small blue bars). Parent SY138 (½ green ½ red) that participated in Crosses_2, _3, and _4, contributed to Crosses_2 and _3, but less to Cross_4 (as seen on parental contribution, Fig. 7.2B).

Results of an analysis with pooled microsatellite data from the four crosses (Fig. 7.2C) showed genetic distinctness of progeny of each cross, *i.e.* the progeny of Cross_1 was almost uniformly red, that of Cross_3 blue and that of Cross_4 yellow. If one of the parents in a cross did not participate in any of the other crosses, then the genetic diversity of the offspring was higher (Cross_1, Cross_3 and Cross_4) than when both of the parents took part in other crosses (Cross_2).

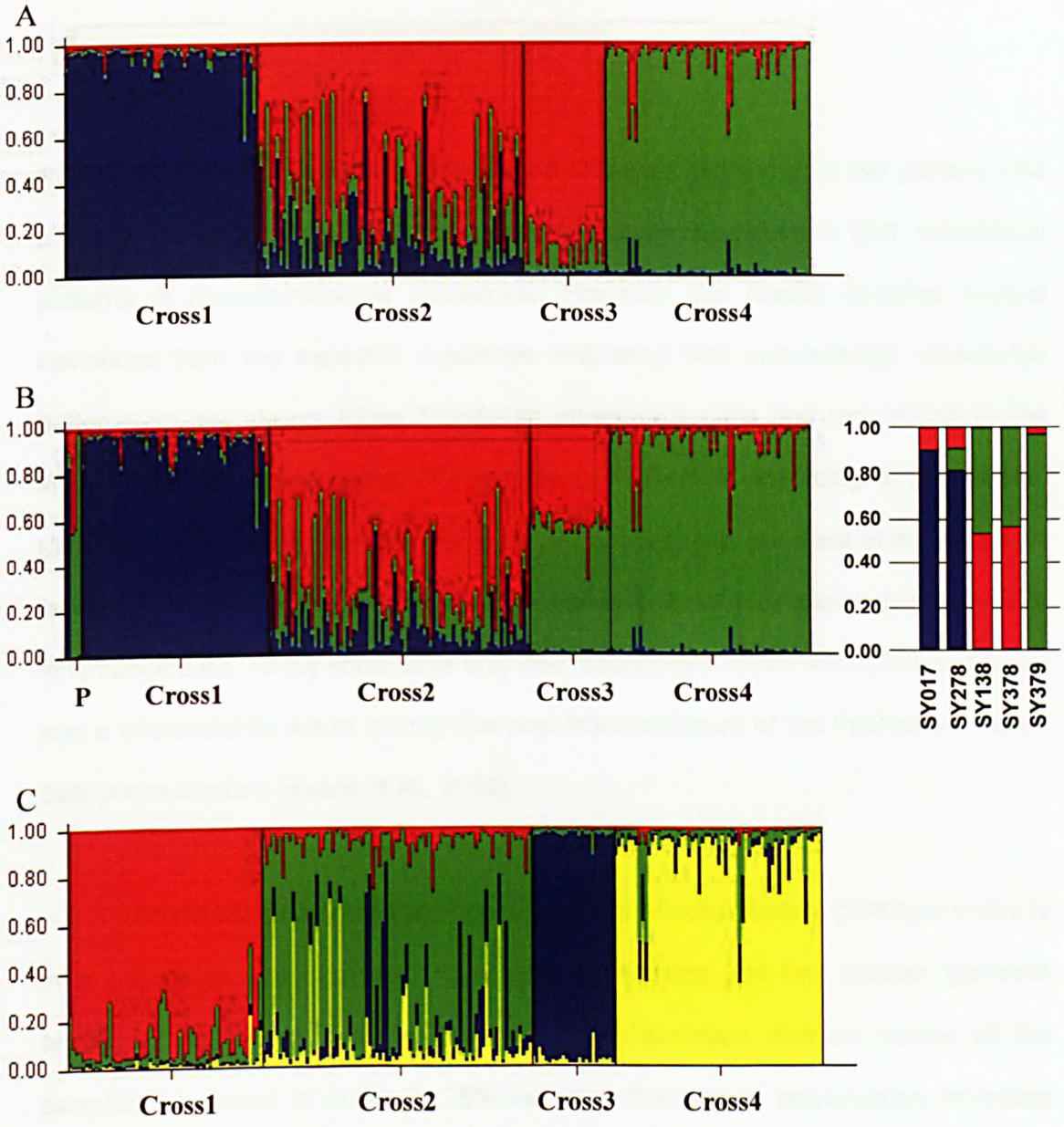


Figure 7.2: Estimated genetic population structure among parental (P) and F1 genotype over the four crosses; Black lines separate the genotypes of parents and those of each of the F1 of the four crosses. Common parameter settings, using STRUCTURE, are: microsatellite data without missing values (N=231), length of burn-in 10000, MCMC repeats 50000 and Admixture models. Strains used were F1 only for Figs. 7.2A and 7.2C and both F1 and parental strains for Fig. 7.2B. 'Allele frequencies' was set as correlated (7.2A and 7.2B) and independent (7.2C). A zoom on parental strains on Fig. 7.2B has been realised to visualise the parental contribution. The parental genotypes are from the left to the right SY017, SY278, SY138, SY378 and SY379 .

7.4. Discussion

Results of the microsatellite genotype and ITS-type screening of the parents and progeny clarified several issues regarding sexual reproduction and inheritance patterns in *Pseudo-nitzschia multistriata*. However, the results revealed several deviations from the expected outcomes, indicating that microsatellite inheritance patterns do not always follow Mendelian inheritance rules and regulations to the letter, and that, in rare cases, ITS regions do not inherit according to predictions. Many explanations can be given for such deviant patterns, but most of those can be falsified in the present study because of the availability of both microsatellite and ITS inheritance data. To my knowledge only one recent paper deals with ITS-1 and ITS-2 and a microsatellite set to identify the population structure of the freshwater diatom *Sellaphora capitata* (Evans et al., 2009).

The results clearly illustrate how sexual reproduction boosts genotype diversity in a population. As expected, four crosses involving just five parents generate already a plethora of different genotypes. The genotype diversity across all the sampled F1 strains of about 70.25% indicates that sexual reproduction increases genotypic diversity.

Sexual reproduction is possible between strains independent of population assignment, and the progeny was equally fecund whether the parents were assigned by STRUCTURE to the same or different populations. Apparently there are no pre- or postzygotic barriers between strains assigned to different populations, and neither

are there any prezygotic ones given the readiness with which the parents in Cross_1 and Cross_2 mated. Of course, tested strains are pampered in lab cultures under conditions very different from those in the field. And in the mating experiments no choice is offered among different suitors, and the possibility of different timings of the sex events in the two populations is not even explored here. Future testing for partner choice is feasible by offering a test-strain the choice between two partner strains of different cell sizes and then screening the resulting pairs of gametangia for bias in partner choice.

The observation, that the two crosses between parents assigned to different populations were more vigorous than those between partners of the same population, does not permit any generalisation, such as inter-population crosses are more vigorous than intra-population ones, because only four types of crosses between pairs of only five different parents were conducted. Many other factors can explain the differences. Crosses_1 and _2 were carried out with parents from different populations, but both involved strain SY017, the partners were of smallish cell size; Instead, Crosses_3 and _4 were carried out with parents assigned to the same population and involving at least one strain of relatively tall partner cells. Taller cells may be more reticent towards mating because of being closer to the upper size limit (D'Alelio et al., 2010) for sexual reproduction.

The genetic identity of strains generated from sister-cells produced on the same gametangium is in accordance with the fact that after the first meiotic division one nucleus goes pycnotic and disappears (Round et al., 1990) while during the second meiotic division, the two chomatids of the chromosomes in the remaining nucleus split at their centromeres, and the resulting haploid sister chromosomes

segregate during anaphase over two haploid gametes per gametangium. The genotypically identical sister-cells are not identical twins from the same zygote because the sister-cells arise from independent fusion of two genotypically identical male gametes with two genotypically identical female gametes, with genotypic identity referring to the seven loci assessed. The result also suggests that recombination through crossing-over between the parental chromosomes of the sexualized cells is not frequent enough to cause a deviation from genetic identity in the sister-cells pair of zygotes, and consequently, the pair of initial cells. Yet, this statement needs back-up by examining other such cases. Unfortunately, catching sister-cells emerging from the pair of auxospores on the same gametangium is no sinecure; timing is of the essence and one has to be sure that one catches exclusively the sister-cells.

The genetic identity of strains generated from sister-cells produced on the same gametangium cannot result from auto-auxosporulation, a process recently demonstrated in *P. brasiliana* (Quijano-Scheggia *et al.*, 2009). Such a process is feasible if a mutation occurs in one of the strains, which permits a sex change and hence, auto-compatibility of that strain. Casteleyn *et al.* (2009) observed such a sex inversion in *P. pungens*. An argument against auto-auxosporulation could be that the parental strains on their own did not form any auxospores, but some substance exuded by the other strain might be needed to induce auto-auxosporulation (Casteleyn et al., 2009). Nonetheless, these explanations are falsified because in the case of auto-auxosporulation the progeny should have exhibited only the genotype of one of the parents, or a recombinant genotype with input from only one of the parents. Yet, the sister-cells CRD1 and CRD2 have inherited alleles from both of the

parents. In addition, the sister-cells exhibited an ITS A/B-type indicating that the ITS A-type and B-type parents have added equally to the genomes of their offspring.

The present results corroborate the observation in D'Alelio *et al.* (2009a) that sexual reproduction success does not depend on the ITS-type of the parental strains because in Crosses_1 and _2, strains of the ITS A-type and those of the ITS B-type mated without fuss. The observation that crosses involving ITS A-type and ITS B-type parents generate almost always progeny of the ITS A/B-type in which ITS-A and ITS-B are more or less equally presented in the electropherograms suggests that the progeny inherits equal amounts of these different copies from their parents. Nonetheless, a few exceptions were observed. The two F1 strains of the ITS B-type among the progeny of Cross_3 (ITS-A/B x ITS-B) can result from coincidental segregation of a string of ITS B-type cistrons in a gamete (Griffiths et al., 1993) resulting from meiosis of an ITS-A/B nucleus and fusion of this gamete with an ITS B-type gamete from the ITS B-type parent. The occurrence of three (out of 34) F1 strains exhibiting the ITS A-type among the progeny from Cross_1 (ITS-A x ITS-B) is harder to explain. These ITS A-types were F1 strains, composed of tall cells. Moreover, they exhibit microsatellite genotypes as expected for F1 strains originating from both parents, thus falsifying the possibility of auto-auxosporulation or strain auto-compatibility. A possible explanation is gene conversion (Stacey, 1994) in which the whole string of rDNA cistrons inherited from the ITS B-type parent is replaced, through gene conversion by an integral copy of the string of rDNA cistrons inherited from the ITS A-type parent. But then it is somewhat hard to explain why it happened the same way in three independent F1 strains.

In any case, the observation that parents of the ITS A/B-type can generate progeny of a A- or B-type or A/B A-dominant or B-dominant type adds a biological explanation for the fact that such ITS-types persist in the samples of *P. multistriata* at the LTER-MC in the Gulf of Naples. However, it cannot be the only explanation for the sudden increase of the proportion of ITS A-types in the sampling campaign of 2009 out of nowhere. Immigration of such types from elsewhere is needed.

As expected, microsatellite alleles generally inherit according to Mendelian rules (Hancock, 1999) and parental genotypes generally contribute equally to the genotypes of the offspring. Of course, Mendelian rules apply with the restriction that during meiosis half of the genetic material goes pycnotic and is eliminated (Round et al., 1990). In a few cases, conspicuous deviations from inheritance according to Mendelian rules were observed. Most of the observed deviations from expectation can be explained away because of scoring difficulties, but that is a technical difficulty. Appearance of rare novel alleles is probably due to the addition or deletion of core repeats in the microsatelite resulting from replication error in a cell (Hancock, 1999) after the strain was genotyped but before it was allowed to reproduce sexually, or due to unequal crossing-over during meiosis (Griffiths et al., 1993).

The observation of an F1 strain that possesses only paternal alleles on one of its loci (*e.g. PNm1*, Cross_2) could be explained *ad hoc* assuming gene conversion in the zygote or emerging initial cell. In that case, a copy of the allele from the paternal chromosome replaces the maternal allele on the chromosome inherited from the mother. Alternatively, a mutation may have happened in one of the two primer annealing sites of the maternal locus, thus rendering that allele not longer amplifiable (null allele). Homozygosity for allele 117 on locus *PNm1* (Cross_2: SY017 x SY138)

can not be a null allele. Expectation is that the strain is heterozygous 117 127 for the same locus. If a null allele occurs, it should be located on SY017 and also be presented in the progeny of the Cross_1. No evidence of null allele occurrence has been detected on the parental genotype SY017 over Cross_1, nor was it demonstrated in Crosses_3 and _4 considering the other parental strain, SY138.

The F1 strain (CS272A2) that apparently inherited both alleles (*PNm7*, *PNm6*, *PNm16* for Cross_4) from their maternal strain is less easily explained away *ad hoc*. One *ad hoc* explanation needs three independent mutations on three loci that change the alleles on these loci conform to those on the maternal strain, followed by propagation of these changes in a cell line within the paternal strain. In that case, the strain would include a minor proportion of cells with the alternative genotypic make-up. The difficulty with this *ad hoc* explanation is that mutations are needed on three independent loci within one and the same cell line within a strain and that these changes generate alleles conform to those in the future partner strain. An alternative *ad hoc* explanation is that the changes may have happened in the progeny after the crosses took place. Gene conversion can now at least explain why the changes are similar to alleles inherited from the maternal strain, but the *ad hoc* hypothesis now needs to explain away why gene conversion performs the same replacements on three independent loci in the same F1 cells, knowing that no linkage evidence were provided between those loci.

Linkage dis-equilibrium was observed for the two loci: *PNm2* and *PNm16*. There is linkage dis-equilibrium if the two loci tested are physically linked (same chromosome) or functionally linked (Devlin & Risch, 1995; Reich & Goldstein, 1999). Possible deviation from Mendelian inheritance was monitored on *PNm16* but not on *PNm2*. Therefore this hypothesis must be rejected.

The one strain in the progeny of Cross_4 that showed a deviant genotype (AA instead of Aa) could be explained away assuming a somatic mutation propagating mitotically and through random drift through part of the parental strain before they had sex, or assuming a mutation during meiosis. Both can give rise to this deviant F1 genotype. However, the deviation was seen on three loci in the same F1 strain, meaning that the particular cell-line within this strain must have accumulated three mutations independently. The deviant genotypes in the progeny of Cross_1 on one locus (observed 100% AA and 0% Aa; expected 50/50 AA and Aa) might be due to random drift of a mutation through all of the culture after this parent was genotyped, but then the same pattern should also show up in other crosses with SY017, which it does not. The deviant genotypes might also be due to the occurrence of a null allele on SY278. However, the parental strain SY278 showed the two alleles when screened together with the F1.

In some cases, F1 strains exhibited not two, but three different alleles. *Pseudo-nitzschia multistriata* is a diploid species with normally two alleles per locus. One explanation is a somatic mutation and propagation of that change through random drift through part of the strain. To find if this is the case, several subcultures

need to be established from single cells isolated from the strain, which need then to be screened for the locus showing three alleles in the original strain. If all the subcultures show the same three alleles persisting on that locus then more *ad hoc* explanations are needed such as failure of the chromatids to separate during the second meiotic division and subsequent migration of a supernumerary chromosome into one of the gametes (Griffiths et al., 1993). This phenomenon may explain cases of unequal parental contribution to the offspring and the various small variations from Mendelian law monitored over crosses.

The low isolation success of F1 cells from Cross_3 (8.25%) might be due to a particular fragility of initial cells resulting from this cross. However, isolation of initial cells from the other crosses was not particularly successful either. Initial cells are tall and relatively fragile; micropipetting them out of a culture plate and washing them through a row of sterile seawater droplets was apparently too stressful for many. The strains that survived the procedure grew as well as any other strains in the collection of progeny. All attempts were performed within a few days. The particularly low isolation success of Cross_3 coincides with one of the parents being of the ITS A/B-type, but there was also just one such type of cross.

Technical troubles were also encountered in the sequencing of the ITS regions of the 253 strains because only 122 were analysed successfully for their ITS-types pattern. Failures were due to reading problems, bad electropherograms and other technical mishaps. Note that the success rate of the microsatellite reading was far higher (246 of 253). Therefore, the problems were not due to quality of the DNA used for the PCR reactions because microsatellite amplification reactions were generally more sensitive to impaired DNA quality than those of ITS, as ITS regions are multi-copy markers.

7.5. Conclusions

Sexual reproduction gives rise to viable progeny irrespective of parental ITS-types and microsatellite population assignment.

Microsatellite inheritance follows Mendelian inheritance rules in the vast majority of cases.

The identical genotypes of two sister cells isolated from the same gametangium supports the idea that one of the nuclei formed after the first meiotic division is eliminated.

Sex increases the genotypic diversity in a population. If the four crosses are considered as representative of the whole *P. multistriata* field bloom, the mean diversity after a sex event may be lower (70.25%) than excepted.

CHAPTER EIGHT

INHERITANCE OF INTERNAL TRANSCRIBED SPACERS AND MICROSATELLITE ALLELES DURING VEGETATIVE REPRODUCTION

8.1. Introduction

Inheritance patterns of internal transcribed spacers (ITS) of the rDNA cistrons and of microsatellite markers were assessed during sexual reproduction in Chapter 7. Therefore, several strains were screened to select a set of parents with microsatellite genotypes that suited the experiments conducted in Chapter 7 best. Just to permit direct comparison of the genotypes of the parental strains and their offspring, the genotypes of the parents were obtained again at the time when the crossing experiments were conducted, *i.e.*, after several months of culture. Genotypic differences were observed in many of the strains. The aim of the present chapter is to explore these differences in order to obtain insight in substitution rates of the microsatellite loci and of the ITS regions during mitotic reproduction in monoclonal cultures of *Pseudo-nitzschia multistriata*.

Living organisms can replicate their DNA into two identical copies, and so can viruses, using the replication machinery of their hosts. Cells divide mitotically into genetically identical daughter cells, and so on, to form a clonal colony (genet) (Collado-vides, 2002). In *Pseudo-nitzschia*, sister-cells hold on to one another at their valve apices by means of their raphe system, thus forming the stepped chains characteristic for this genus (Chapter 1). Cells in such chains are thus genetically identical copies.

In theory, mitosis generates exact daughter copies of the mother nucleus. In practice changes can happen during DNA-replication such as single point mutations, insertions, deletions, *etc.* Some regions in the genome (*e.g.* hot spots regions such as SNP, microsatellite or region enriched in GC) have a capacity to change faster than null recombination regions (*e.g.* centromere and telomere) (Drouaud *et al.*, 2006). Mutations can be neutral (*i.e.* not affecting gene functions) adaptive/beneficial (*i.e.* increasing organism fitness) or deleterious. Mechanisms exist that repair these errors afterwards, but in many cases there is no way for a repair mechanism, such as gene conversion, to recognize the original state from the changed state.

On microsatellites, polymerase slippage leads to the addition or deletion of short repeats in the core region (rate: $3.6x10^{-6}$ for CA cores (Shinde *et al.*, 2003)). Changes on microsatellite loci are not exactly random events, either. The frequency of polymerase slippage is positively related to the number of repeats and negatively related to the length of these individual repeats in a microsatellite (Amos 2009).

Repair mechanisms are not impartial either. Following replication, when the double helix is re-established, the length of the microsatellite alleles on the homologous chromosomes may differ, either because of additions/deletions of simple core repeats during DNA replication, or because of differences inherited via the parents (that is also a reason why they could be different). In those cases, DNA heteroduplexes can be recognised and changed, especially through homologous gene conversion (Barbera & Petes 2006).

Very few studies have dealt with inheritance of genetic markers during asexual reproduction; in unicellular organism studies are recorded on full length viral RNA sequences (Miralles *et al.*, 1999), mutated genes in bacteria (Arjan *et al.*, 1999),

mitochondrial gene and 18S sequence regions in microalgae (Colegrave, 2002; Grimsley *et al.*, 2010). To my knowledge, this is the first study addressing the variability of microsatellite markers over vegetative cell lines in a diatom.

8.2. Materials and Methods

Twenty-seven strains of *Pseudo-nitzschia multistriata* were selected randomly from the early-summer and autumn blooms of both year 2008 and 2009 (Table 8.1). Protocols for cell isolation, culturing, DNA extraction and long term maintenance of cultures were presented in Chapters 2 and 3. The DNA of these strains was extracted twice. The first extraction (E1) was conducted on a culture of fast-growing cells maintained at 22 °C, a few weeks after strain establishment. The second extraction (E2) was performed on the strains after 3 to 16 months of incubation under conditions permitting slow growth (15 °C; see Chapter 3 for further details of maintenance). The DNA of one strain (SY017) was extracted three times: the first extraction (E1) a few weeks after strain establishment, the second ($E2_{ITS}$) and the third extraction (E3) after 7 and 16 months of culture at 15 °C, respectively (Table 8.1).

ITS regions (Chapter 4) and seven microsatellite loci (Chapter 5) have been used to uncover somatic mutations that occurred during the months of clonal growth and that were swept by random drift to dominate the culture at the time of the second DNA extraction. Twenty-four of the 27 strains were screened for ITS variability (Table 8.1; methods, see Chapter 4) and 26 of the 27 strains were screened subsequently for their microsatellite pattern (Table 8.2, method described in Chapters 5 and 6) over

the two extractions mentioned above. ITS-types (Table 8.1) and microsatellite genotypes (Table 8.2) were compared per strain over the two extractions.

For microsatellite genotyping, the probability of observing a substituted allele per locus over the 26 strains (*i.e.* p(*1_substitution/locus_26 strains*)) was calculated as the sum of single allele substitution over the total number of loci ($N_{LOCI\ SUBSTITUTED}$) minus loci presenting null alleles ($N_{NULL\ LOCI}$) (Table 8.3):

- p(*1_substitution/locus_26 strains*) = 26 strains [($N_{LOCI\ SUBSTITUTED} - N_{NULL\ LOCI}$) x 2 alleles]).

The expected number of double-changed loci was calculated as the probability of observing the two substituted alleles per locus over the 26 strains tested p(*2_substitutions/locus_26 strains*):

- p(*2_substitutions/locus_26 strains*) = p(*1_substitution/locus_26 strains*)^2.

Growth experiments were conducted over four strains (SY413, SY416, SY486 and SY487, Chapter 3) incubated at 15 °C. The growth rate was calculated during the exponential phase of the culture (Table 3.1 in Chapter 3); however over one month of culture (*ca.* 31 days), strains generally undergo a short exponential phase (*ca.* one week) and a long plateau phase (*ca.* three weeks, Chapter 3) assuming that during the plateau the growth rate is close to zero. To calculate the growth rate over a month of culture, the exponential and plateau phases were integrated (Table 8.4). The one-month growth period was estimated from a semi-log plot of cell concentration (cells·ml^{-1}) over time (Chapter 3). Least-square regression was applied

to the selected log (base 10) data. Logarithmic transformation from log base 10 to log base 2 was used to estimate the number of division per day (Guillard, 1973).

Taking into account that exponential growth rates differ among strains, the mean growth rate (0.070 division·day^{-1}) and the Upper/Lower growth rates (0.047 and 0.108 division·day^{-1}) were used to calculate the mean and its limits of the number of mitotic divisions over that period and the substitution rate per strain (Table 8.5).

The Number of mitotic divisions (N_G) and the substitution rate (R_{SUB}) were estimated as follow:

- $N_G = N_{MONTHS}$ x Growth rate over month at 15°C
- $R_{SUB} = N_{SUBSTITUTION} / N_G$

BioEdit (version 7.0.9.0, (Hall et al., 1992)) was used to identify ITS-types. CEQ^{TM}2000XL (version 4.3.9, Beckman Coulter^{TM}) was used to infer microsatellite length. Microsatellites lengths were classed in length-windows determined during the stabilisation of protocol (Chapter 5). Genetix (version 4.05.2) was applied to screen for linkage dis-equilibrium.

8.3. Results

No changes were observed on the ITS-types in any of the strains screened at the two extraction times (*i.e.* SY017, three extraction points over time). Instead, several alleles on the microsatellite loci showed changes (Table 8.2). A total of 32 loci over the 182 possible loci (7 microsatellite markers x 26 strains) exhibited changes. Apart from that, one locus showed the presence of three alleles and one locus revealed potential null alleles, but those were not counted.

With regard to individual alleles, 40 alleles over the 364 allele entries present in the data set had been substituted: 24 alleles (6.59% of the data base) showed one substitution per locus, and, eight loci (4.39% of the data base) presented a substitution on both alleles (Table 8.3). Changes were observed on: *PNm2* (7 alleles), *PNm5* (14), *PNm3* (2), *PNm6* (6), *PNm16* (1) and *PNm1* (10). Only *PNm7* exhibited no changes among any of the strains screened. The number of stepwise substitution per allele ranged from 0 to 7 and the substitution rate per allele ranged from 0 to 0.28 (Table 8.3).

The substitution rate per locus was correlated positively with the number of core repeats of the alleles on that locus (Fig. 8.1, Table 8.3). The outlyier to the trend was *PNm5*, which had about the same core length as *PNm6*, but showed a far higher substitution rate ($p_{SUBST./LOCUS}$= 0.56 and 0.23, respectively). Moreover, the substitution rate per locus was higher in di-nucleotides (*PNm2*, *PNm5*, *PNm6*, and *PNm1*) than in tri-nucleotides (*PNm3* and *PNm16*; Fig. 8.1).

Table 8.1: Characteristics of strains from which genomic DNA was extracted two or three times (zero to 16 months apart; N_{MONTHS}). 'Bloom period' refers to from which blooming period the strain was established; Cell size refers to the length over the apical axis (in µm) at the time of strain establishment. The two or three extraction dates (E1, E2, E3) for ITS-type and microsatellite marker screening. E_{ITS} and E_{MSAT} indicate extraction and subsequent screening of only the indicated marker). ITS-type (A, B and A/B). A 'n.t.' refers to not tested.

Strain code	Bloom period	Cell size (µm)	DNA extraction dates	N_{MONTHS}	ITS-type
SY016	2008 early-summer	48	$E1_{ITS}$: 01/10/2008 $E2_{ITS}$: 01/10/2008	0	B
SY017	2008 early-summer	48	E1: 01/10/2008 $E2_{ITS}$: 13/05/2009 E3: 19/01/2010	7 16	A A
SY042	2008 autumn	45	E1: 08/10/2008 E2: 15/02/2010	16	B
SY077	2008 autumn	77	E1: 15/10/2008 E2: 29/01/2010	15	A/B
SY081	2008 autumn	42	E1: 07/11/2008 E2: 29/01/2010	15	A/B
SY138	2008 autumn	36	E1: 11/11/2008 E2: 19/01/2010	14	B
SY174	2008 autumn	44	E1: 13/05/2009 E2: 29/01/2010	9	A/B
SY244	2008 autumn	41	E1: 18/02/2009 E2: 29/01/2010	11	A/B
SY341	2009 early-summer	56	E1: 18/09/2009 E2: 29/01/2010	4	A/B
SY368	2009 early-summer	50	E1: 28/07/2009 E2: 02/02/2010	6	A/B
SY373	2009 early-summer	45	E1: 28/07/2009 E2: 02/02/2010	6	A
SY375	2009 early-summer	53	E1: 28/07/2009 E2: 02/02/2010	6	B
SY378	2009 early-summer	48	E1: 28/07/2009 E2: 02/02/2010	6	A/B
SY379	2009 early-summer	50	E1: 18/08/2009 E2: 19/01/2010	5	A
SY380	2009 early-summer	49	E1: 18/08/2009 E2: 09/02/2010	6	A

Table 8.1: Continued.

SY384	2009 early-summer	47	E1: 18/08/2009 E2: 29/01/2010	5	B
SY386	2009 early-summer	44	E1: 18/08/2009 E2: 02/02/2010	6	A/B
SY392	2009 early-summer	49	E1: 18/08/2009 E2: 02/02/2010	6	A/B
SY394	2009 early-summer	49	E1: 18/08/2009 E2: 02/02/2010	6	B
SY399	2009 early-summer	46	E1: 18/08/2009 E2: 02/02/2010	6	B
SY400	2009 early-summer	45	E1: 18/08/2009 E2: 04/02/2010	6	B
SY407	2009 early-summer	44	E1: 18/08/2009 E2: 04/02/2010	6	A/B
SY410	2009 early-summer	n.t.	E1: 18/08/2009 E2: 29/01/2010	5	A/B
SY411	2009 early-summer	n.t.	E1: 18/08/2009 E2: 29/01/2010	5	A/B
SY416	2009 autumn	42	$E1_{MSAT}$: 14/10/2009 $E2_{MSAT}$: 26/01/2010	n.t.	n.t.
SY481	2009 autumn	49	$E1_{MSAT}$: 20/10/2009 $E2_{MSAT}$: 29/01/2010	n.t.	n.t.
SY576	2009 autumn	42	$E1_{MSAT}$: 13/10/2009 $E2_{MSAT}$: 29/01/2010	n.t.	n.t.

Table 8.2: Microsatellite genotypes of the 26 strains tested on the genomic DNA isolated at time E1 and at time E2 (dates of E1 and E2 for that strain see Table 8.1). N_M denotes number of months between subsequent DNA extractions. The figures in the matrix denote microsatellite lengths of the two alleles for each of the seven loci tested: PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1. Figures in orange indicate alleles that present a single substitution on a locus for that strain; green figures denote loci on which both alleles have changed the same way.

Strain	N_M	E	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
SY416	3	E1	175	187	238	238	258	262	208	208	249	255	337	337	119	127
		E2	175	187	238	238	258	262	208	208	249	255	337	337	117	127
SY481	3	E1	175	187	238	238	258	258	208	208	255	267	319	319	117	119
		E2	175	187	236	236	258	258	208	208	255	267	319	319	117	119
SY576	4	E1	175	187	238	238	258	258	208	208	255	271	319	337	117	127
		E2	175	185	236	238	258	258	208	208	255	271	319	337	117	127
SY341	4	E1	175	185	238	238	262	262	208	208	247	255	319	325	117	117
		E2	175	187	238	238	262	262	208	208	247	255	319	325	117	119
SY379	5	E1	175	175	230	238	262	262	208	208	265	265	319	325	117	125
		E2	175	175	230	238	262	262	208	208	267	267	319	325	117	127
SY384	5	E1	175	187	238	238	258	258	208	208	255	271	319	337	117	127
		E2	175	187	236	238	258	258	208	208	255	271	319	337	117	127
SY410	5	E1	175	187	230	238	258	262	208	208	255	255	319	337	115	127
		E2	175	187	230	236	258	262	208	208	255	255	319	337	115	127
SY411	5	E1	175	187	238	238	258	262	208	208	249	255	319	325 337	117	127
		E2	175	187	236	238	258	262	208	208	249	249	319	337	117	127
SY368	6	E1	175	175	238	242	258	262	208	208	255	255	319	325	115	127
		E2	175	175	238	242	258	262	208	208	255	255	319	325	117	127
SY373	6	E1	175	187	238	238	262	262	208	208	249	255	319	325	115	115
		E2	175	187	238	238	262	262	208	208	249	255	319	325	117	117
SY375	6	E1	175	187	238	240	258	262	208	208	249	267	307	319	115	127
		E2	175	187	236	238	258	262	208	208	249	267	307	319	115	127
SY378	6	E1	175	187	236	238	262	262	208	208	249	249	325	337	117	117
		E2	175	187	238	238	262	262	208	208	249	249	325	337	115	117
SY380	6	E1	175	187	238	238	258	262	208	208	249	255	307	319	117	117
		E2	175	187	238	238	258	262	208	208	249	255	307	319	115	119
SY386	6	E1	175	187	230	238	258	262	208	208	255	267	307	319	117	127
		E2	175	185	230	238	258	262	208	208	255	267	307	319	117	127

Table 8.2: Continued.

Strain	N_M	E	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
SY392	6	E1	175	187	230	238	262	262	208	208	249	255	325	337	117	147
		E2	175	187	230	**236**	262	262	208	208	249	255	325	337	117	147
SY394	6	E1	175	187	238	238	258	258	208	208	255	271	319	337	117	127
		E2	175	**185**	**236**	238	258	258	208	208	255	271	319	337	117	127
SY399	6	E1	175	187	238	238	258	258	208	208	255	271	319	337	115	127
		E2	175	187	238	238	258	258	**205**	**205**	255	271	319	337	**117**	127
SY400	6	E1	175	187	238	238	258	258	208	208	255	271	319	337	117	127
		E2	175	187	238	238	258	258	208	208	255	271	319	337	117	127
SY407	6	E1	175	187	238	242	262	262	208	208	255	267	319	337	115	127
		E2	175	187	238	**240**	262	262	208	208	255	267	319	337	**117**	127
SY174	9	E1	175	175	238	240	258	262	205	208	255	261	325	337	127	129
		E2	175	175	238	**242**	258	262	205	208	255	261	325	337	127	**127**
SY244	11	E1	175	175	236	240	258	262	205	208	255	261	325	337	127	127
		E2	175	175	236	240	258	262	205	208	255	261	325	337	127	127
SY138	14	E1	181	185	238	238	258	258	208	208	255	255	307	337	117	127
		E2	**175**	**187**	238	238	258	258	208	208	255	255	307	337	117	127
SY077	15	E1	175	175	238	242	258	262	205	208	255	261	325	337	127	127
		E2	175	175	**236**	**240**	258	262	205	208	255	261	325	337	127	127
SY081	15	E1	175	175	0	0	258	262	205	208	253	259	325	337	127	127
		E2	175	175	236	240	258	262	205	208	**255**	**261**	325	337	127	127
SY042	16	E1	175	187	238	238	258	262	205	208	255	267	325	337	117	117
		E2	175	**175**	238	**242**	258	262	205	208	255	**261**	325	337	117	117
SY017	16	E1	175	175	238	242	262	262	205	208	261	267	325	337	127	127
		E3	175	175	238	242	262	262	205	208	261	267	325	337	127	127

A total of 40 clear substitutions were observed over the data set. The probability of an entry changing was 0.110 (*i.e.* 26 strains; 7 loci; p(*1_mutation/locus_26 strains*)=0.110). The probability that changes hit the same entry twice was of 0.0121 (*i.e.* 26 strains; 7 loci; p(*2_mutations/locus_26 strains*)=0.0121), which is virtually the same as the probability of a hit at an allele and a hit at the second allele in the same locus. The expected number of loci exhibiting a double-change was 2.2 (0.0121 x 7 loci x 26 strains), whereas the observed value was eight (*i.e.* 4 homozygote-loci and 4 heterozygote-loci). Double-mutated loci were present at all tested loci, with the exception of *PNm7* and *PNm16* (Table 8.2). All changes were characterised by the addition/deletion of one core repeat (di-nucleotides core for *PNm2, PNm5, PNm6* and *PNm1* loci and tri-nucleotides core for *PNm3*).

Homozygosity could potentially be over-scored due to the presence of null alleles. The presence of null alleles was possible in *PNm3* (SY399), *PNm5* (SY481), and *PNm1* (SY373) but was infirmed for locus *PNm6* (SY379, Cross_4 in Chapter 7).

Linkage dis-equilibrium may influence the mutation rate; however, none of the strains showed repeated and common substitution patterns or associated patterns of changes for homozygote or heterozygote double-substituted loci (Table 8.2). Even if, results of GENETIX showed potential linkage dis-equilibrium between locus pairs *PNm2-PNm16* and *PNm2-PNm1*, *PNm3-PNm6* and *PNm3-PNm16*, and *PNm6-PNm16*.

Table 8.3: Number of substitutions (N_{SUB}), number of Pseudo-nitzschia multistriata *strains tested (N_{STR}), the substitution rate per allele ($p_{SUB./AL}$) and per locus ($p_{SUB./LOC}$) over the 14 alleles within the 7 loci (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1). The second row indicates the microsatellite core repetition (di- or tri-nucleotides) and the respective fluorophore colours (CY5 (blue) and IRD700 (green)) used in microsatellite length detection.*

Locus	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
Core	$(AC)_{17}$		$(GT)_{11}$		$(CA)_8$		$(GAC)_8$		$(CT)_9$		$(GTC)_7$		$(TC)_{23}$	
N_{SUB}	1	6	7	7	0	0	1	1	2	4	0	1	7	3
N_{STR}	26	26	25	25	26	26	26	26	26	26	26	26	26	26
$P_{SUB./AL}$	0.04	0.23	0.28	0.28	0.00	0.00	0.04	0.04	0.08	0.15	0.00	0.04	0.27	0.12
$P_{SUB./LOC}$	0.27		0.56		0.00		0.08		0.23		0.04		0.38	

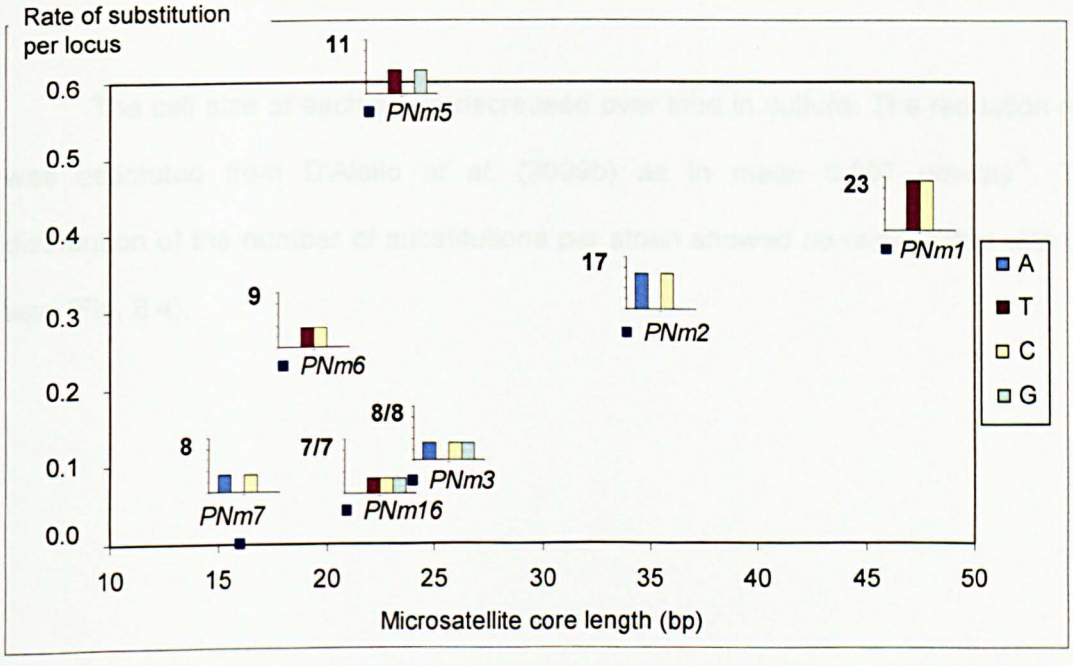


Figure 8.1: Rate of mutation per microsatellite markers (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1) as function of microsatellite core length (in base pairs) and the proportion of A (blue), T (red), C (yellow) and G (light blue) microsatellite core content. Bold numbers indicate the expected number of C or G in the microsatellite sequence. Two and three columns indicate respectively di- and tri-nucleotides core.

The number of substitutions per strain ranged between 0 and 3, and did not show a relationship with the estimated number of mitotic divisions between the two measurements (Table 8.5, Fig. 8.2). The mutation substitution rate per mitotic division had a tendency to decrease over time. Several strains did not show any substitution (*e.g.* at SY400 (12 mitotic divisions), SY244 (24) and SY017 (33), Table 5) whereas others showed substitutions after only few mitotic divisions (*e.g.* SY481 at 7 mitotic divisions). The general trend indicated that after *ca.* 27 mitotic divisions the substitution rate was lower than 5.10^{-2} while after 42 mitotic divisions the substitution rate approached to zero (Fig. 8.3).

The cell size of each strain decreased over time in culture. The reduction rate was estimated from D'Alelio *et al.* (2009b) as in mean 0.007 $\mu m \cdot day^{-1}$. The distribution of the number of substitutions per strain showed no relationship with cell size (Fig. 8.4).

Table 8.4: Estimated growth rate of four strains (SY413, SY416, SY486 and SY487) over ca. one month of incubation at 15 °C. Columns report the code of the strains tested, the length of the incubation time (in days), the growth rate over the incubation time (in divisions per day).

Strain code	Incubation time (days)	Growth rate (divisions·day^{-1})
SY413	33	0.049
SY416	33	0.108
SY486	33	0.076
SY487	29	0.047
Mean	32	0.070 (± 0.029)

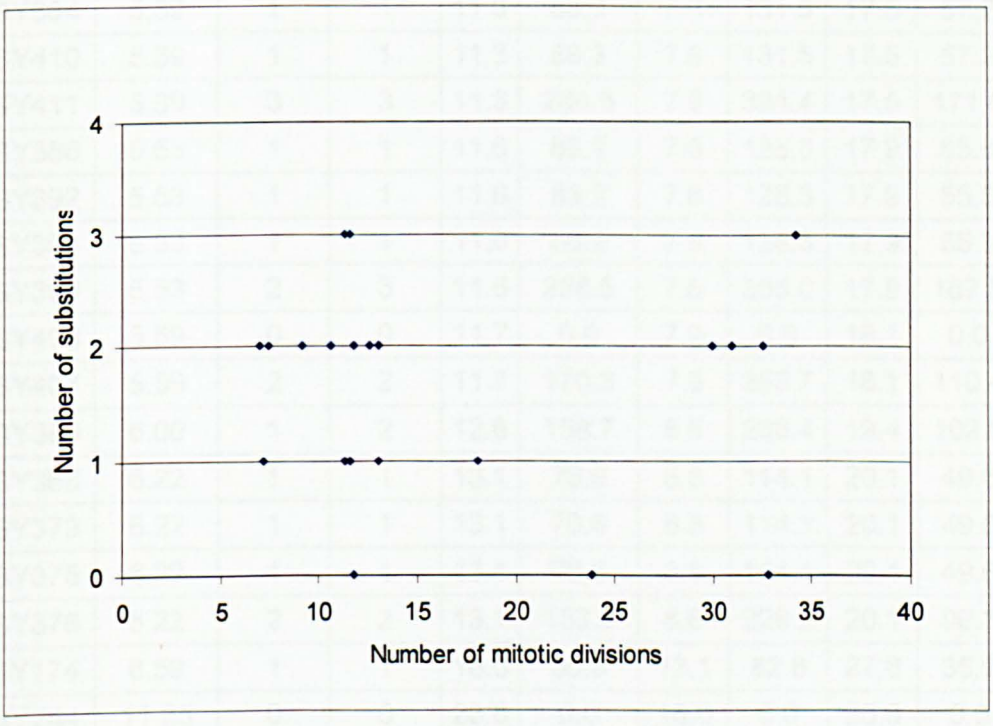


Figure 8.2: Number of substitutions over the number of estimated cell divisions between the subsequent DNA samplings in the 26 strains of Pseudo-nitzschia multistriata *tested for microsatellites markers.*

Table 8.5: Estimation of the mutation rate over the number of mitotic divisions for each the 26 strains tested with microsatellite markers; the strain code, the number of months between two DNA-extractions (N_{MONTHS}), the number of loci showing substitutions ($N_{LOCUS\ SUB}$), the number of substituted alleles ($N_{ALLELE\ SUB}$), the number of mitotic divisions (N_G) and the allelic substitution rate (R_{SUB}). Mean UPPER and LOWER limits stands for mean estimates of N_G and R_{MUT} considering the different growth rate (Table 8.4).

Strain code	N_{MONTHS}	$N_{LOCUS\ SUB}$	$N_{ALLELE\ SUB}$	TENDENCY		UPPER limit		LOWER limit	
				N_G	R_{SUB} (10^{-3})	N_G	R_{SUB} (10^{-3})	N_G	R_{SUB} (10^{-3})
SY481	3.32	1	2	7.0	286.7	4.7	426.9	10.8	185.8
SY416	3.42	1	1	7.2	139.2	4.8	207.3	11.1	90.2
SY576	3.55	2	2	7.5	268.1	5.0	399.3	11.5	173.8
SY341	4.38	2	2	9.2	217.7	6.2	324.2	14.2	141.1
SY379	5.07	1	2	10.6	188.0	7.1	280.0	16.4	121.9
SY384	5.39	1	1	11.3	88.3	7.6	131.5	17.5	57.2
SY410	5.39	1	1	11.3	88.3	7.6	131.5	17.5	57.2
SY411	5.39	3	3	11.3	264.8	7.6	394.4	17.5	171.6
SY386	5.53	1	1	11.6	86.2	7.8	128.3	17.9	55.8
SY392	5.53	1	1	11.6	86.2	7.8	128.3	17.9	55.8
SY394	5.53	1	1	11.6	86.2	7.8	128.3	17.9	55.8
SY399	5.53	2	3	11.6	258.5	7.8	385.0	17.9	167.5
SY400	5.59	0	0	11.7	0.0	7.9	0.0	18.1	0.0
SY407	5.59	2	2	11.7	170.3	7.9	253.7	18.1	110.4
SY380	6.00	1	2	12.6	158.7	8.5	236.4	19.4	102.9
SY368	6.22	1	1	13.1	76.6	8.8	114.1	20.1	49.6
SY373	6.22	1	1	13.1	76.6	8.8	114.1	20.1	49.6
SY375	6.22	1	1	13.1	76.6	8.8	114.1	20.1	49.6
SY378	6.22	2	2	13.1	153.2	8.8	228.2	20.1	99.3
SY174	8.59	1	1	18.0	55.5	12.1	82.6	27.8	35.9
SY244	11.35	0	0	23.8	0.0	16.0	0.0	36.8	0.0
SY138	14.28	1	2	30.0	66.7	20.1	99.4	46.3	43.2
SY081	14.74	1	2	30.9	64.6	20.8	96.3	47.7	41.9
SY077	15.49	1	2	32.5	61.5	21.8	91.6	50.2	39.8
SY017	15.63	0	0	32.8	0.0	22.0	0.0	50.6	0.0
SY042	16.28	3	3	34.2	87.7	23.0	130.7	52.8	56.9

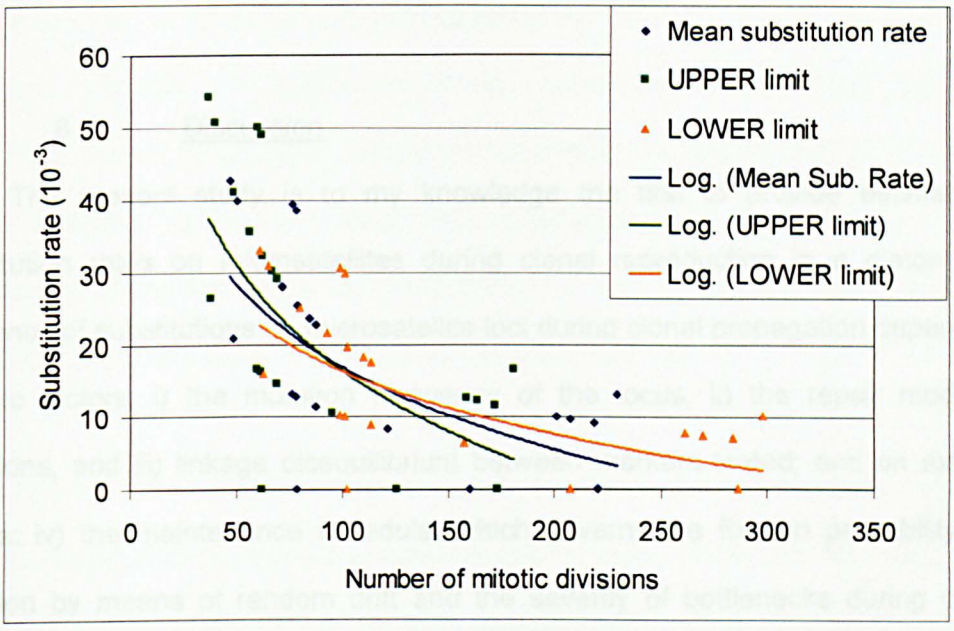


Figure 8.3: Substitution rates over number of mitotic divisions of the Pseudo-nitzschia multistriata *strains. The mean substitution rates over the number of mitotic divisions (blue diamonds) and the upper (green rectangles) and lower (orange triangles) limits have been estimated as explained in the text. Mean substitution rate: y=-110.89Ln(x)+412.6, Upper limit: y=-165.15Ln(x)+548.73 and Lower limit: y=-71.873Ln(x)+298.6), with respective colours.*

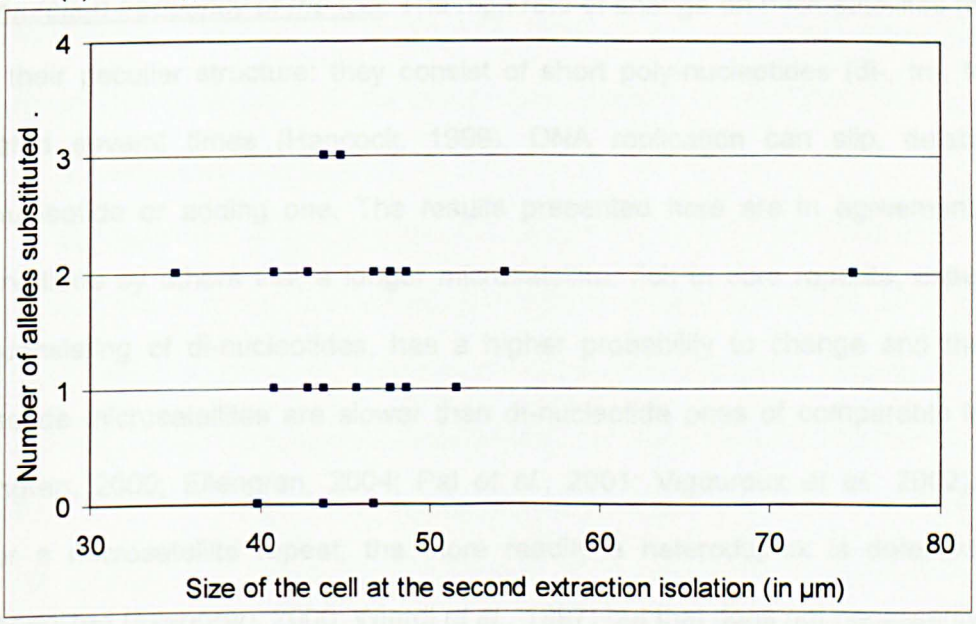


Figure 8.4: : Number of substitutions recorded in the 26 strains of Pseudo-nitzschia multistriata *as a function of their cell size after the second extraction for microsatellites markers test. The cell size was estimated from the size reduction rate in cultures from previous studies on* P. multistriata *(0.007 µm·day^{-1}, D'Alelio* et al. *2009b).*

8.4. Discussion

The present study is to my knowledge the first to provide estimates of substitution rates on microsatellites during clonal reproduction in a diatom. The frequency of substitutions on microsatellite loci during clonal propagation depends on intrinsic factors: i) the mutation frequency of the locus, ii) the repair modes of mutations, and iii) linkage disequilibrium between markers tested; and on extrinsic factors: iv) the maintenance schedule, which governs the fixation probability of a mutation by means of random drift and the severity of bottlenecks during culture transfer, and v) whether the locus is selectively neutral or not under the conditions at which the strain is maintained.

The mutation frequency of the loci: The high rate of change on microsatellites results from their peculiar structure: they consist of short poly-nucleotides (di-, tri-, tetra-) repeated several times (Hancock, 1999). DNA replication can slip, deleting a polynucleotide or adding one. The results presented here are in agreement with observations by others that a longer microsatellite, rich in core repeats, especially one consisting of di-nucleotides, has a higher probability to change and that tri-nucleotide microsatellites are slower than di-nucleotide ones of comparable length (Ellengren, 2000; Ellengren, 2004; Pal *et al.*, 2001; Vigouroux *et al.*, 2002). The longer a microsatellite repeat, the more readily a heteroduplex is detected and homogenized (Ellengren, 2000; Wierdl *et al.*, 1997) but that does not necessarily lead to the elimination of a change, it can also lead to the incorporation of the change in the other chromosome as well.

According to Amos (2009) microsatellites exhibit a "life cycle" with a tendency for alleles to increase in length until hitting a species-dependent upper threshold. Nearer the threshold, repair is biased more towards the shorter copies to "prevent infinite expansion" (Ellengren, 2000). Yet, results show no skew among the changes towards shortening the longer microsatellites. The reason might be that the longest microsatellites in this study are not hitting the threshold, yet.

The absence of any changes observed in the ITS regions over extended periods of clonal division are expected for two reasons. First, the ITS regions are multi-copy markers, rendering changes on individual copies undetectable in the electropherograms of the PCR products, even if the clonal progeny of the cell with that particular change on one of the ITS copies is going to dominate the strain. All the other ITS copies on the other cistrons still swamp the change. In contrast, microsatellites are co-dominant single copy markers and a mutation on one of the two sites in the diploid genome of a diatom does become visible if the offspring of the mutated cell takes over the strain.

In the case of *P. multistriata*, the ITS regions contain microsatellite-like repeats as well, and two of these microsatellites exhibit a variable number of cores among the strains tested. But even so, changes are not expected to be as frequent as those on the seven regular microsatellites explored in this study because the microsatellite core repeats on the ITS are short, and the cores themselves are often relatively long (core of two to five nucleotides; Chapter 4), in comparison to those in the microsatellites used in this study (Fig. 8.1). Notably, the ITS microsatellite with the shortest core-repeat (TA) was also the most polymorphic in the collection of strains, causing stutter in the electropherograms (Chapter 4, D'Alelio *et al.* 2009a).

The number of substitutions recorded in the data set of *P. multistriata* (26 strains) is far higher than expected given observations of the rates of change in other organismal groups. In yeast, common somatic mutations recorded are base-substitutions ($3.3.10^{-9}$), microsatellite mutations ($7.2.10^{-7}$ for 15 core repeats to $7.2.10^{-5}$ for 36 core repeats (Lynch *et al.*, 2008)) and recombination events (crossing-over at 4.10^{-5} per division and non-reciprocal recombination rate at $1.2.10^{-5}$ per division in yeast (Barbera & Petes, 2006)).

Results show no correlation between the substitution frequency and cell size. Size can be viewed as a proxy of age of the clone because in diatoms, older cell lines consist of smaller cells (MacDonald 1869, Pfitzer 1869). But this comparison is not entirely applicable to our strains because the cell line has not spent all of its lifetime in a culture.

The repair modes of the mutations at a given locus: DNA repair and non-crossing-over recombination events (*e.g.* gene conversion) can homogenise the emerging differences but there is no way to recognize the change from the original state. Many changes in the data set are of the type where a heterozygous locus becomes a homozygous locus. These changes may come about through non-crossing-over recombination (*i.e.* exchange of regions among paired chromosomes by means of gene conversion), which occur not only during meiosis (Storlazzi *et al.*, 1995) but also during mitosis (Barbera & Petes, 2006). Alternatively, the differences maybe explained by the emergence of null alleles

The maintenance schedule: Culture maintenance affects the fixation probability of a mutation through random drift by means of the severity of bottlenecks during culture transfer. It is assumed that random drift sweeps individual mutations to dominance. Drift is probably fairly efficient because the long-term maintenance schedule is comparable to a population going through a monthly cycle of a 'boom' followed by an extreme bottleneck in which only about 76 cells are transferred into new medium. Most mutations are likely to be lost, but a few 'fortunate' ones take part in the next 'boom,' thus increasing their proportion dramatically before entering the next transfer-bottleneck.

Colegrave (2002) demonstrates that the inoculum size linearly affected the rate of adaptation of the culture undergoing asexual reproduction. The reason may be that the more the number of cells in a strain, and thus the more the carriers of mutations make through the transfer bottleneck and therefore, the more the genotypic variation enters the new culture on which natural selection can work. However, there is no direct evidence for such a phenomenon in the cultures maintained.

Natural selection: The decrease in the substitution rate with the number of mitotic divisions (Fig. 8.3) is peculiar, and it is contrary to finding by Miralles *et al.* (1999) that older viral strains exhibit higher mutation rates. An explanation for the decrease in substitution rate with age of culture could be that the microsatellite loci are not entirely selectively neutral, or they are located in the vicinity of non-neutral markers. Recent studies on unicellular organisms such as viruses (Miralles *et al.* 1999), *Escherichia coli* (Arjan *et al.* 1999) and the microalga *Chlamydomonas reinhardtii*

(Colegrave 2002) confirm that multiple beneficial mutations can occur in mono-clonal cultures undergoing asexual reproduction. The higher the number of beneficial mutations inherited asexually (Colegrave 2002, Miralles *et al.* 1999), the higher the probability that those mutations are transmitted and sexually inherited, even if DNA-repairing mechanisms and recombination events occur.

The initially extremely high substitution rates and the decrease in the rate over the mitotic generations in culture may result from strains adapting to the stable environment of the culture (Metzgar & Wills, 2000; Sniegowski *et al.*, 2000). Conditions in a clonal laboratory culture differ radically from those experienced in the field. These lab conditions might lead initially to highly increased substitution rates driven by selection for faster growth in culture. With time progressing, the strain may gain adaptedness to the new situation, and the fixation rate may settle down to the normal background value (Sniegowski et al., 2000).

8.5. Conclusions

This Chapter shows that mitotic substitutions accumulated during ongoing mitotic cell division. Substitutions do happen in the microsatellite markers tested and their rates of change depend on the core composition (numbers of nucleotides, core repeats and G/C content), on the strain tested (growth rate, genetic adaptation, *etc.*) and on the presence of potential null alleles. Highest mutation rates were recorded for highly repeated di-nucleotides. Mutations were however independent from the cell size of the strain, the number of generations and the linkage disequilibrium between microsatellites loci.

A peculiar result was the non-linear decrease of substitutions in microsatellite markers with increasing number of cell divisions (months in culture) due to selection

and adaptation pressure. An explanation is that the transfer from the field to the laboratory conditions generates an initial flurry of fixations of mutations, mutations that would never stand a chance in the field, but that are advantageous in the lab culture conditions. The subsequent decrease could then result from the strain settling down in the new culture regime. Microsatellite markers are a good tool to answer this question. However, a monitoring of a few species over different mitotic divisions is needed to assess the fixation rate and to understand the role of substitutions (*e.g.* are they deleterious?).

CHAPTER NINE

GENERAL CONCLUSIONS

The planktonic diatom *Pseudo-nitzschia multistriata* was observed at the long term ecological research station MareChiara (LTER-MC) in the Gulf of Naples (GoN) for the first time in 1995 (Sarno & Dahlman 2000) and has bloomed there ever since in the autumn, and more recently (2004) also in the early-summer. This thesis focuses on population genetic structure of *P. multistriata* species using the Internal Transcribed Spacers (ITS) localised in the ribosomal DNA as well as seven polymorphic microsatellites. Cells were isolated from plankton samples taken in 2008 and 2009 and grown into strains. From these strains ITS-sequences and microsatellite fingerprints were gathered.

For this work it was needed to maintain strains over extended periods of time to permit gathering different kinds of data as the need arises. However, fast vegetative division leads to a rapid cell size reduction (Macdonald 1869, Pfitzer 1869) and therefore to a reduction of the time a strain can be maintained in culture. The method of slow-growth used in the culture maintenance of *P. multistriata* over extended periods has been developed step-by-step using previous publications on *P. multistriata* culture conditions (Churro *et al.* 2009; D'Alelio *et al.* 2009a, 2009b; Orsini *et al.* 2002; Rhodes *et al.* 2000). This technique ensures the survival of several hundreds of strains for population genetic research over several months to a few years, and provides material for DNA extraction and crossing experiments.

In a modelling exercise of cell size distributions in the cell counts at LTER-MC, D'Alelio *et al.* (2010) inferred that the species exhibits a biennial life cycle with a new cohort appearing once every two years and sex occurring only in the fall of the second. In another study, D'Alelio *et al.* (2009a) observed variation in the ITS-regions of rDNA of *P. multistriata* in the GoN. However, those differences did not constitute a barrier to sexual reproduction. Thus, unlike findings in other morphological species in *Pseudo-nitzschia* (Amato *et al.* 2006), the ITS differences were no signature of cryptic species diversity in *P. multistriata*. Moreover, D'Alelio *et al.* (2009a) observed in their strain collection a gradual decrease in the proportion of ITS A-type and B-type and an increase in the A/B-type, suggesting a merger between two populations of different ITS-type.

The goals of the present thesis were to determine the genetic structure of the *P. multistriata* population(s) over two years, one year during which sexual reproduction occurred and one year of strictly vegetative reproduction, in the Gulf of Naples. Specifically, ITS-screening was set out to test the hypothesis of merging-populations one of ITS-A and one of ITS-B in the Gulf of Naples, proposed by D'Alelio *et al.* (2009a). ITS-type distributions over strains sampled in 2008 and 2009 refute this hypothesis. Indeed, the proportion of ITS-pure types increased markedly in 2009, results that were not expected from the hypothesis of merging populations.

Population structure of *P. multistriata* over the 2008-2009 blooms in the GoN showed two distinct populations (F_{ST} = 0.091), each of them being in Hardy Weinberg equilibrium, and an important number of migrants (N_m=3.014) between them. ITS-type proportions were also different between the two populations further supporting

their biological distinctness. This indicates that population differentiation can occur in plankton species in a region, and these different populations can occur in sympatry. The disappearance of one of the two 2008-populations in 2009 suggests high dynamism and flush-out during the winter, which is somewhat counterintuitive given the existence of distinct populations. ITS-type proportions within the 2009-population were comparable between the early-summer and the autumn sampling period indicating that the population remains present throughout the growing season (at least in 2009).

Crossing experiments showed that strains from either population were sexually compatible, suggesting that there are no pre- or postzygotic barriers between the two populations. A possible explanation is that the two populations may live side-by-side, but if the timing of their sexual reproduction is slightly offset, then they can remain genetically distinct. Future work should find methods to detect sexual events of this species in the field and screen those who reproduce and those who do not, for their population assignment. Another interesting perspective is geographic sampling to check for geographical differentiation. *P. multistriata* has been monitored over the Mediterranean Sea in Spain, Italy and Greece (Quijano-Scheggia *et al.* 2008a, 2008b; Sarno & Dahlman 2000; Moschandreou & Nikolaidis 2010). If all ITS variants are similar and sexual compatibility is proven between the three localities, then the species *P. multistriata* is the same over the Mediterranean Sea. However, if the microsatellite loci uncover distinct populations over the three localities this may indicate that populations are isolated. However, if some strains match with other localities this will imply that *P. multistriata* is influenced by the Mediterranean circulation.

Sexual reproduction increased genotypic diversity. The genotypic diversity over crosses ranged between 43.48% and 90.91%. This genotype diversity was independent of the ITS-types because F1 strains of the same genotype belonged to different ITS-types. Every cross seems to generate a new genetically distinct sub-population; however, crosses that share one mate showed genetic similarities. By extrapolation to the field where mating is random, clonal dispersal is high and the probability of inter-crosses elevated, the various groups in the parental cohort will merge through sex into an offspring cohort that should be genetically closely related to, but not genetically the same as the parental groups.

Several deviations were observed from Mendelian inheritance rules for microsatellites. Seven of the 28 loci tested over the four crosses presented such deviations due to the presence of new alleles (step-wise mutation) or inheritance dis-equilibrium (meiosis separation). The progeny, over three loci tested, inherited 50% of each parental genetic pool, at the exception of Cross_4 where inheritance dis-equilibrium favour the female gene pool. Genotypic changes not only occur during sexual reproduction events, but also during vegetative reproduction (Barbera *et al.* 2006). Several substitutions were recorded in a series of strains re-screened after several months of maintenance. Microsatellites that presented a di-nucleotide core evolved faster than tri-nucleotide ones and the longer microsatellites with di-nucleotide cores evolved more frequently than shorter ones.

Under asexual reproduction, microsatellite loci appear to mutate independently of the number of generations undergone or the cell size. An explanation could be that

culture transfer progressively selects the strains (bottleneck, Colegrave 2002) or that the strains are selected by adaptation (Amos, 2009; Metzgar & Wills, 2000; Sniegowski *et al.*, 2000). With selection, the number of potential mutations per strain decreases. Unfortunately, in our experiment mutation fixation over the same group of strains was not assessed over time but analysis of strains was performed after observing microsatellite patterns slightly differing between the two screens. Nevertheless, the results show the occurrence of somatic changes in microsatellite loci during asexual reproduction.

A similar study of *P. multistriata* population genetic structure over the summer and autumn bloom of 2010 is now on-going. If no radical flush-out happened in the winter of 2010, then a similar microsatellite genotype as 2009 should be observed such as a new cohort, composed of tall cells and arising from the sexual event of 2009. In the same way, the ITS patterns of the small cells should be the same as the one visible in 2009, and the ITS pattern of the new cohort should show a higher proportion of ITS A/B-mixed types. Any radical deviation from that expectation suggests immigration from genetically distinct populations from elsewhere, which should be detectable with microsatellite data.

APPENDICES

Appendix 1: Recapitulative tables of strains sampled at MareChiara (MC) station from July 2008 to November 2009 (across two early-summer blooms (☼) and two autumnal blooms). First column indicates the bloom period during which strains were sampled. Second column indicates the code of each strain. Third column indicates the ITS-types identify as ITS-A (blue), ITS-B (red) and ITS-A/B types. NR indicates that the ITS-types were not successfully identified. Blast-search indicated that the culture was contaminated. The (-) indicates missing values. Fourth column indicates the genetic population identified by microsatellite markers (Msat pop) as population 1 (green), 2 (orange) and intermediate strains (IS). Fifth column indicates the size of each strain after isolation (in micrometer). The columns six and seven indicate respectively the code and the date of MareChiara sampling.

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2008 ☼	SY014	A/B	-	60	823	29/07/08
	SY016	B	-	48	823	29/07/08
	SY017	A	-	48	823	29/07/08
2008 autumn bloom	SY018	A/B	1	42	827	28/08/08
	SY019	NR	1	-	827	28/08/08
	SY020	A	-	-	828	02/09/08
	SY021	A/B	IS	45	828	02/09/08
	SY025	A/B	-	48	830	17/09/08
	SY027	A/B	1	48	830	17/09/08
	SY028	NR	2	48	830	17/09/08
	SY031	B	1	42	830	17/09/08
	SY033	NR	-	45	830	17/09/08
	SY034	A/B	1	39	830	17/09/08
	SY035	NR	IS	60	830	17/09/08
	SY038	A/B	1	45	830	17/09/08
	SY039	A/B	IS	48	830	17/09/08
	SY040	NR	1	45	830	17/09/08
	SY041	NR	1	42	830	17/09/08
	SY042	B	2	45	830	17/09/08
	SY043	NR	2	45	830	17/09/08
	SY044	A/B	1	42	830	17/09/08
	SY045	A/B	1	42	830	17/09/08
	SY046	B	2	42	830	17/09/08
	SY047	NR	1	48	830	17/09/08
	SY048	NR	1	42	830	17/09/08
	SY049	A/B	1	42	830	17/09/08
	SY050	A/B	1	42	830	17/09/08
	SY051	NR	-	48	830	17/09/08
	SY053	A/B	-	42	830	17/09/08
	SY054	B	-	42	830	17/09/08
	SY056	A/B	1	-	830	17/09/08
	SY057	A/B	1	-	830	17/09/08
	SY058	A/B	1	42	830	17/09/08
	SY061	A/B	-	42	831	23/09/08
	SY062	A/B	1	39	831	23/09/08
	SY064	A/B	-	42	831	23/09/08
	SY065	NR	-	42	831	23/09/08
	SY067	A/B*	1	45	831	23/09/08
	SY068	NR	1	48	831	23/09/08
	SY070	A/B	-	42	831	23/09/08
	SY071	A/B*	1	42	831	23/09/08

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2008 autumn bloom	SY073	-	-	42	831	23/09/08
	SY074	A/B	IS	48	831	23/09/08
	SY075	A/B*	-	45	831	23/09/08
	SY077	A/B	1	77	831	23/09/08
	SY079	A/B	2	42	831	23/09/08
	SY080	A/B	1	42	831	23/09/08
	SY081	A/B	-	42	831	23/09/08
	SY082	A/B	1	41	831	23/09/08
	SY083	A/B	1	48	831	23/09/08
	SY085	A/B	1	36	831	23/09/08
	SY086	A/B	2	-	831	23/09/08
	SY087	A/B	1	45	831	23/09/08
	SY088	A/B	1	-	831	23/09/08
	SY089	A/B	-	45	831	23/09/08
	SY090	B	2	42	831	23/09/08
	SY091	A/B	-	36	831	23/09/08
	SY095	A/B	-	42	831	23/09/08
	SY097	A/B	1	45	831	23/09/08
	SY098	A/B	2	45	831	23/09/08
	SY099	A/B	1	42	831	23/09/08
	SY100	B	IS	42	831	23/09/08
	SY101	A/B	1	42	831	23/09/08
	SY103	NR	1	42	831	23/09/08
	SY104	A/B	-	42	831	23/09/08
	SY105	A/B	1	42	831	23/09/08
	SY106	A/B	-	42	831	23/09/08
	SY108	A/B	1	42	831	23/09/08
	SY109	A/B	1	42	831	23/09/08
	SY110	NR	2	42	831	23/09/08
	SY112	NR	1	45	831	23/09/08
	SY113	B	2	42	831	23/09/08
	SY114	A/B	1	39	831	23/09/08
	SY115	NR	2	42	831	23/09/08
	SY116	A/B	1	42	831	23/09/08
	SY117	A/B	1	42	831	23/09/08
	SY118	A/B	1	42	831	23/09/08
	SY119	NR	1	42	831	23/09/08
	SY120	A/B	2	48	831	23/09/08
	SY122	A/B*	1	45	831	23/09/08
	SY123	A/B	-	42	831	23/09/08
	SY125	-	-	42	831	23/09/08

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2008 autumn bloom	SY126	A/B	-	42	831	23/09/08
	SY127	A/B	IS	42	831	23/09/08
	SY130	A/B	1	-	831	23/09/08
	SY131	A/B	1	45	831	23/09/08
	SY132	A/B	1	42	831	23/09/08
	SY133	NR	1	42	831	23/09/08
	SY134	NR	1	42	831	23/09/08
	SY136	A/B	2	48	831	23/09/08
	SY137	NR	1	42	831	23/09/08
	SY138	B	2	36	831	23/09/08
	SY140	A/B	-	42	831	23/09/08
	SY141	A/B	-	39	831	23/09/08
	SY142	A/B	1	42	831	23/09/08
	SY143	NR	-	42	831	23/09/08
	SY144	NR	-	45	831	23/09/08
	SY145	A/B	1	42	831	23/09/08
	SY146	A/B	1	42	831	23/09/08
	SY147	A/B	1	43	831	23/09/08
	SY148	A/B	1	44	832	23/09/08
	SY149	A/B	1	43	832	23/09/08
	SY151	A/B	1	43	832	23/09/08
	SY152	A/B	1	42	831	23/09/08
	SY153	NR	1	42	831	23/09/08
	SY154	A/B	-	42	831	23/09/08
	SY155	A/B	2	42	831	23/09/08
	SY156	A/B	1	42	831	23/09/08
	SY157	A/B	1	42	831	23/09/08
	SY159	A/B	-	46	831	23/09/08
	SY160	A/B	1	43	832	30/09/08
	SY161	A/B	1	42	832	30/09/08
	SY162	A/B	2	44	832	30/09/08
	SY167	A/B	1	41	832	30/09/08
	SY168	A/B	1	41	832	30/09/08
	SY169	A/B	1	43	832	30/09/08
	SY170	A/B	1	43	832	30/09/08
	SY171	A/B	2	49	832	30/09/08
	SY172	A/B	1	43	832	30/09/08
	SY173	A/B	1	48	832	30/09/08
	SY174	A/B	1	44	832	30/09/08
	SY177	-	-	43	832	30/09/08
	SY179	B	2	59	832	30/09/08

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2008 autumn bloom	SY182	A/B	1	42	832	30/09/08
	SY183	A/B	1	44	832	30/09/08
	SY184	B	2	41	832	30/09/08
	SY185	A/B	1	42	832	30/09/08
	SY186	A/B	1	44	832	30/09/08
	SY188	NR	IS	42	832	30/09/08
	SY190	A/B	1	38	832	30/09/08
	SY191	A/B	-	43	832	30/09/08
	SY192	A/B	1	43	832	30/09/08
	SY193	A/B	-	40	832	30/09/08
	SY195	A/B	1	43	832	30/09/08
	SY196	A/B	2	42	832	30/09/08
	SY197	A/B	-	43	832	30/09/08
	SY199	A/B	1	42	832	30/09/08
	SY200	A/B	1	42	832	30/09/08
	SY201	A/B	1	42	832	30/09/08
	SY202	A/B	1	42	832	30/09/08
	SY203	A/B	1	43	832	30/09/08
	SY206	A/B	1	44	832	30/09/08
	SY207	A/B	1	46	832	30/09/08
	SY208	A/B	1	44	832	30/09/08
	SY209	A/B	1	44	832	30/09/08
	SY210	-	-	44	832	30/09/08
	SY211	A/B	-	46	832	30/09/08
	SY212	A/B	-	46	832	30/09/08
	SY213	B	2	42	832	30/09/08
	SY214	A/B	1	45	832	30/09/08
	SY216	A/B	1	42	832	30/09/08
	SY217	A/B	1	45	832	30/09/08
	SY218	A/B	1	46	832	30/09/08
	SY219	A/B	1	46	832	30/09/08
	SY221	A/B	1	43	832	30/09/08
	SY222	A/B	1	-	838	11/11/08
	SY226	A/B	1	44	832	30/09/08
	SY227	A/B	1	44	832	30/09/08
	SY228	A/B	-	47	832	30/09/08
	SY231	A/B	1	40	832	30/09/08
	SY234	A/B	1	42	835	21/11/08
	SY237	B	2	41	834	14/10/08
	SY238	A/B	1	42	834	14/10/08
	SY239	A/B	1	41	834	14/10/08

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2008 autumn bloom	SY240	NR	1	-	834	14/10/08
	SY241	A/B	1	42	834	14/10/08
	SY243	A/B	IS	41	834	14/10/08
	SY244	A/B	1	41	834	14/10/08
	SY246	A/B	1	39	834	14/10/08
	SY247	A/B	-	41	834	14/10/08
	SY248	NR	1	42	834	14/10/08
	SY249	B	2	41	834	14/10/08
	SY250	B	2	42	834	14/10/08
	SY251	A/B	1	43	834	14/10/08
	SY252	B	2	42	834	14/10/08
	SY253	B	2	41	834	14/10/08
	SY254	A/B	1	45	834	14/10/08
	SY255	A/B	1	42	834	14/10/08
	SY256	A/B	1	44	834	14/10/08
	SY259	A/B	1	43	834	14/10/08
	SY274	A/B	-	43	835	21/11/08
	SY278	B	2	40	837	04/11/08
	SY279	A/B	1	42	837	04/11/08
	SY280	A/B	1	43	837	04/11/08
	SY281	A/B	1	43	837	04/11/08
	SY282	A/B	2	37	837	04/11/08
	SY285	A/B	-	42	837	04/11/08
	SY288	A/B	1	43	837	04/11/08
	SY289	B	2	46	837	04/11/08
	SY290	A/B	1	41	838	11/11/08
	SY292	A/B	1	41	838	11/11/08
	SY294	B	-	54	838	11/11/08
	SY296	-	-	43	838	11/11/08
	SY297	NR	1	41	838	11/11/08
	SY299	B	2	40	838	11/11/08
	SY300	A/B	1	40	838	11/11/08
	SY301	A/B	1	43	838	11/11/08
	SY302	A/B	1	41	839	18/11/08
	SY304	B	2	54	839	18/11/08
	SY305	A/B	1	36	839	18/11/08
	SY306	A/B	1	41	839	18/11/08
	SY307	A/B	1	36	839	18/11/08
	SY308	A/B	1	43	839	18/11/08
	SY309	A/B	2	37	839	18/11/08
	SY310	A/B	1	42	842	09/12/08
	SY314	A/B	-	42	851	17/02/09

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 early summer bloom	SY317	A/B*	2	50	868	30/06/2009
	SY318	A	2	54	868	30/06/2009
	SY319	B	2	48	868	30/06/2009
	SY320	NR	-	46	868	30/06/2009
	SY321	NR	2	50	868	30/06/2009
	SY322	NR	2	54	868	30/06/2009
	SY323	NR	-	49	868	30/06/2009
	SY324	A/B	2	50	868	30/06/2009
	SY325	A/B	2	50	868	30/06/2009
	SY326	NR	2	49	868	30/06/2009
	SY327	A/B	2	48	868	30/06/2009
	SY328	A/B	2	54	868	30/06/2009
	SY329	A	2	50	868	30/06/2009
	SY330	A/B	2	50	868	30/06/2009
	SY331	NR	2	48	868	30/06/2009
	SY332	NR	-	49	868	30/06/2009
	SY333	A/B	2	49	868	30/06/2009
	SY334	NR	2	47	868	30/06/2009
	SY335	NR	2	48	868	30/06/2009
	SY336	A	2	48	868	30/06/2009
	SY337	A/B	2	48	868	30/06/2009
	SY338	A	2	47	868	30/06/2009
	SY339	A/B	2	47	868	30/06/2009
	SY340	B	-	54	868	30/06/2009
	SY341	A/B	2	56	868	30/06/2009
	SY342	A	2	43	868	30/06/2009
	SY343	A/B	2	50	868	30/06/2009
	SY344	A/B	2	48	868	30/06/2009
	SY345	NR	-	43	868	30/06/2009
	SY346	NR	2	48	869	07/07/2009
	SY347	NR	2	50	869	07/07/2009
	SY348	B	2	49	869	07/07/2009
	SY349	NR	2	52	869	07/07/2009
	SY350	NR	2	50	869	07/07/2009
	SY351	NR	-	46	869	07/07/2009
	SY352	NR	-	54	869	07/07/2009
	SY353	A/B	2	40	869	07/07/2009
	SY354	NR	2	56	869	07/07/2009
	SY355	NR	2	47	869	07/07/2009
	SY356	NR	2	48	869	07/07/2009
	SY357	NR	2	50	869	07/07/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 early summer bloom	SY358	A/B	2	48	869	07/07/2009
	SY359	A/B	2	43	869	07/07/2009
	SY360	NR	2	49	869	07/07/2009
	SY361	B	2	47	869	07/07/2009
	SY362	B	2	48	869	07/07/2009
	SY363	A/B	-	46	869	07/07/2009
	SY364	NR	2	48	869	07/07/2009
	SY365	NR	2	40	869	07/07/2009
	SY366	NR	2	50	869	07/07/2009
	SY367	B	2	48	869	07/07/2009
	SY368	A/B	2	50	869	07/07/2009
	SY369	NR	2	49	869	07/07/2009
	SY370	NR	-	48	869	07/07/2009
	SY371	B	2	45	869	07/07/2009
	SY372	NR	-	52	869	07/07/2009
	SY373	A/B	2	45	869	07/07/2009
	SY374	NR	2	47	869	07/07/2009
	SY375	B	2	53	869	07/07/2009
	SY376	NR	-	43	869	07/07/2009
	SY377	A/B	-	49	869	07/07/2009
	SY378	A/B*	2	48	869	07/07/2009
	SY379	A	2	50	869	07/07/2009
	SY380	A	2	49	869	07/07/2009
	SY381	NR	2	48	869	07/07/2009
	SY382	B	2	52	869	07/07/2009
	SY383	NR	2	53	869	07/07/2009
	SY384	B	2	47	869	07/07/2009
	SY385	A/B	2	48	869	07/07/2009
	SY386	A/B	2	44	869	07/07/2009
	SY387	A/B	-	48	869	07/07/2009
	SY388	A/B	2	48	869	07/07/2009
	SY389	A/B	2	49	870	14/07/2009
	SY390	NR	-	47	870	14/07/2009
	SY391	NR	2	45	870	14/07/2009
	SY392	A/B	2	49	870	14/07/2009
	SY393	NR	2	53	870	14/07/2009
	SY394	B	2	49	870	14/07/2009
	SY395	B	2	47	870	14/07/2009
	SY396	A/B*	2	55	870	14/07/2009
	SY397	A	IS	51	870	14/07/2009
	SY398	NR	2	46	870	14/07/2009
	SY399	B	2	46	870	14/07/2009
	SY400	B	2	45	870	14/07/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 ☼	SY401	NR	2	46	870	14/07/2009
2009 autumn bloom	SY402	B	2	45	870	14/07/2009
	SY403	NR	2	47	870	14/07/2009
	SY404	A	-	47	870	14/07/2009
	SY405	NR	2	43	870	14/07/2009
	SY406	A/B	2	44	870	14/07/2009
	SY407	A/B	2	44	870	14/07/2009
	SY408	NR	2	46	870	14/07/2009
	SY409	NR	-	47	870	14/07/2009
	SY410	A/B	2	-	871	21/07/2009
	SY411	A/B	-	-	871	21/07/2009
	SY412	NR	2	42	880	22/09/2009
	SY413	A/B	2	42	880	22/09/2009
	SY414	NR	-	41	880	22/09/2009
	SY415	NR	2	54	880	22/09/2009
	SY416	A/B	2	42	880	22/09/2009
	SY417	NR	-	47	880	22/09/2009
	SY418	A/B	2	43	880	22/09/2009
	SY419	B	2	47	880	22/09/2009
	SY420	NR	2	46	880	22/09/2009
	SY421	B	2	54	880	22/09/2009
	SY422	NR	-	48	880	22/09/2009
	SY423	NR	-	47	880	22/09/2009
	SY424	NR	-	45	880	22/09/2009
	SY425	NR	-	44	880	22/09/2009
	SY426	NR	2	42	881	29/09/2009
	SY427	NR	-	42	881	29/09/2009
	SY428	A/B	2	42	881	29/09/2009
	SY429	A/B	-	45	881	29/09/2009
	SY430	NR	2	45	881	29/09/2009
	SY431	A/B	2	44	881	29/09/2009
	SY432	NR	2	43	881	29/09/2009
	SY433	NR	IS	54	881	29/09/2009
	SY434	A/B	2	43	881	29/09/2009
	SY435	A/B	2	43	881	29/09/2009
	SY436	A/B	2	42	881	29/09/2009
	SY437	A/B	2	43	881	29/09/2009
	SY438	NR	2	42	881	29/09/2009
	SY439	NR	2	43	881	29/09/2009
	SY440	A/B	2	42	881	29/09/2009
	SY441	NR	2	42	881	29/09/2009
	SY442	NR	2	41	881	29/09/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 autumn bloom	SY443	NR	2	41	881	29/09/2009
	SY444	NR	IS	47	881	29/09/2009
	SY445	A/B	2	38	881	29/09/2009
	SY446	NR	2	41	881	29/09/2009
	SY447	NR	2	42	881	29/09/2009
	SY448	NR	2	42	881	29/09/2009
	SY449	A/B	2	47	881	29/09/2009
	SY450	B	2	44	881	29/09/2009
	SY451	B	2	41	881	29/09/2009
	SY452	A/B	2	43	881	29/09/2009
	SY453	NR	2	43	881	29/09/2009
	SY454	NR	-	44	881	29/09/2009
	SY455	NR	-	44	881	29/09/2009
	SY456	B	IS	49	881	29/09/2009
	SY457	NR	2	43	881	29/09/2009
	SY458	NR	2	45	881	29/09/2009
	SY459	NR	2	42	881	29/09/2009
	SY460	NR	2	42	881	29/09/2009
	SY461	NR	2	44	881	29/09/2009
	SY462	A/B	-	42	881	29/09/2009
	SY463	NR	2	42	881	29/09/2009
	SY464	NR	2	43	881	29/09/2009
	SY465	B	2	47	881	29/09/2009
	SY466	NR	2	47	881	29/09/2009
	SY467	A/B	2	41	881	29/09/2009
	SY468	NR	2	47	881	29/09/2009
	SY469	NR	2	44	881	29/09/2009
	SY470	NR	2	47	881	29/09/2009
	SY471	NR	2	41	881	29/09/2009
	SY472	NR	2	44	881	29/09/2009
	SY473	NR	2	43	881	29/09/2009
	SY474	NR	2	44	881	29/09/2009
	SY475	NR	2	41	881	29/09/2009
	SY476	NR	2	42	881	29/09/2009
	SY477	NR	2	48	881	29/09/2009
	SY478	NR	2	49	881	29/09/2009
	SY479	NR	2	46	881	29/09/2009
	SY480	NR	2	43	881	29/09/2009
	SY481	A	2	49	881	29/09/2009
	SY482	A/B	2	43	881	29/09/2009
	SY483	NR	IS	52	881	29/09/2009
	SY484	NR	-	40	881	29/09/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 autumn bloom	SY485	NR	2	41	881	29/09/2009
	SY486	A/B	2	40	881	29/09/2009
	SY487	B	2	43	881	29/09/2009
	SY488	NR	2	43	881	29/09/2009
	SY489	NR	-	46	882	06/10/2009
	SY490	NR	2	47	882	06/10/2009
	SY491	NR	2	43	882	06/10/2009
	SY492	NR	-	43	882	06/10/2009
	SY493	NR	-	44	882	06/10/2009
	SY494	NR	2	43	882	06/10/2009
	SY495	NR	-	44	882	06/10/2009
	SY496	NR	-	44	882	06/10/2009
	SY497	NR	-	44	882	06/10/2009
	SY498	NR	-	43	882	06/10/2009
	SY499	NR	2	46	882	06/10/2009
	SY500	NR	-	41	882	06/10/2009
	SY501	NR	2	44	882	06/10/2009
	SY502	NR	-	43	882	06/10/2009
	SY503	NR	-	42	882	06/10/2009
	SY504	NR	-	44	882	06/10/2009
	SY505	A	2	44	882	06/10/2009
	SY506	NR	-	43	882	06/10/2009
	SY507	NR	-	42	882	06/10/2009
	SY508	A	2	46	882	06/10/2009
	SY509	NR	-	42	882	06/10/2009
	SY510	B	2	46	882	06/10/2009
	SY511	NR	-	42	882	06/10/2009
	SY512	NR	-	47	882	06/10/2009
	SY513	NR	2	45	882	06/10/2009
	SY514	B	2	43	882	06/10/2009
	SY515	NR	-	42	882	06/10/2009
	SY516	NR	2	46	882	06/10/2009
	SY517	B	2	43	882	06/10/2009
	SY518	NR	-	41	882	06/10/2009
	SY519	NR	2	44	882	06/10/2009
	SY520	NR	-	43	882	06/10/2009
	SY521	NR	2	44	882	06/10/2009
	SY522	NR	2	43	882	06/10/2009
	SY523	NR	2	43	882	06/10/2009
	SY524	NR	-	42	882	06/10/2009
	SY525	NR	2	42	882	06/10/2009
	SY526	NR	-	42	882	06/10/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 autumn bloom	SY527	A	2	43	882	06/10/2009
	SY528	NR	-	41	882	06/10/2009
	SY529	NR	-	42	882	06/10/2009
	SY530	NR	-	44	882	06/10/2009
	SY531	NR	-	47	882	06/10/2009
	SY532	NR	-	44	882	06/10/2009
	SY533	NR	2	47	882	06/10/2009
	SY534	NR	2	45	882	06/10/2009
	SY535	NR	-	46	882	06/10/2009
	SY536	NR	-	43	882	06/10/2009
	SY537	NR	2	47	882	06/10/2009
	SY538	A	2	42	882	06/10/2009
	SY539	NR	-	48	882	06/10/2009
	SY540	NR	2	44	882	06/10/2009
	SY541	NR	2	43	882	06/10/2009
	SY542	A/B	2	44	883	13/10/2009
	SY543	A/B	2	42	883	13/10/2009
	SY544	NR	2	40	883	13/10/2009
	SY545	NR	2	42	883	13/10/2009
	SY546	NR	-	43	884	20/10/2009
	SY547	NR	2	45	884	20/10/2009
	SY548	NR	2	44	884	20/10/2009
	SY549	NR	-	46	884	20/10/2009
	SY550	NR	-	43	884	20/10/2009
	SY551	NR	2	42	884	20/10/2009
	SY552	NR	-	42	884	20/10/2009
	SY553	NR	2	45	884	20/10/2009
	SY554	NR	2	46	884	20/10/2009
	SY555	NR	-	42	884	20/10/2009
	SY556	NR	-	42	884	20/10/2009
	SY557	NR	-	42	884	20/10/2009
	SY558	NR	2	43	884	20/10/2009
	SY559	NR	-	42	884	20/10/2009
	SY560	NR	-	42	884	20/10/2009
	SY561	NR	2	43	884	20/10/2009
	SY562	NR	2	44	884	20/10/2009
	SY563	NR	2	42	884	20/10/2009
	SY564	NR	2	43	884	20/10/2009
	SY565	NR	-	42	884	20/10/2009
	SY566	NR	2	47	884	20/10/2009
	SY567	NR	2	43	884	20/10/2009
	SY568	B	2	45	884	20/10/2009

Appendix 1. Continued.

Bloom	Strain	ITS-type	Msat pop	Size (μm)	MC code	MC date
2009 autumn bloom	SY569	NR	-	43	884	20/10/09
	SY570	NR	2	43	884	20/10/09
	SY571	NR	-	42	884	20/10/09
	SY572	NR	-	42	884	20/10/09
	SY573	NR	-	41	884	20/10/09
	SY574	NR	-	42	884	20/10/09
	SY575	NR	2	44	885	27/10/09
	SY576	B	2	42	885	27/10/09
	SY577	NR	2	43	885	27/10/09
	SY578	A/B	2	42	885	27/10/09
	SY579	NR	2	42	885	27/10/09
	SY580	A/B	2	44	885	27/10/09
	SY581	A/B	2	42	885	27/10/09
	SY582	B	2	43	885	27/10/09
	SY583	B	2	42	885	27/10/09
	SY584	A/B	2	46	885	27/10/09
	SY585	NR	-	43	885	27/10/09
	SY586	B	2	43	885	27/10/09
	SY587	NR	2	40	885	27/10/09
	SY588	NR	2	44	885	27/10/09
	SY589	NR	2	42	888	17/11/09
	SY590	NR	2	40	888	17/11/09

Appendix 2: Microsatellite patterns of the strains sampled during over early-summer (☼) 2009 and autumnal (✂) blooms of 2008 and 2009 at LTER-MC. The strain code (strain), the allelic comnination over the seven microsatellites tested (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1).

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2008 ✂	SY018	175	175	236	242	258	262	205	208	255	261	325	337	127	127
	SY019	177	177	238	242	258	262	205	208	255	261	325	337	127	127
	SY020	0	0	236	240	258	262	205	208	249	265	301	325	115	127
	SY021	175	175	238	242	258	262	205	208	257	267	325	337	127	127
	SY027	175	175	236	240	258	262	205	208	255	261	337	337	127	127
	SY028	175	175	240	240	258	262	208	208	255	265	307	325	119	119
	SY031	177	187	236	236	258	262	205	208	265	269	325	337	129	129
	SY033	177	177	238	238	258	262	208	208	0	0	325	337	119	119
	SY034	175	175	238	242	258	262	205	208	255	261	325	337	121	127
	SY035	175	187	238	238	258	262	205	205	255	263	319	337	115	127
	SY038	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY039	175	175	238	240	258	262	205	208	267	269	325	337	127	127
	SY040	177	177	238	242	258	262	205	208	255	261	325	337	127	127
	SY041	175	175	236	240	258	262	205	208	255	261	322	337	127	127
	SY042	175	187	238	238	258	262	205	208	255	267	325	337	117	117
	SY043	175	175	238	242	258	262	205	208	255	267	325	337	117	117
	SY044	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY045	175	175	238	242	258	262	205	208	257	263	325	337	129	129
	SY046	175	175	240	242	258	262	205	208	257	269	325	337	117	117
	SY047	175	175	238	242	258	262	205	208	255	261	337	337	127	127
	SY048	177	177	238	242	258	262	205	208	255	261	325	337	127	127
	SY049	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY050	175	175	238	238	258	262	205	208	253	261	301	325	119	119
	SY051	175	175	236	240	258	262	205	208	0	0	325	337	121	127
	SY053	177	177	0	0	258	262	205	208	255	257	325	337	127	127
	SY054	175	187	238	242	0	0	205	208	267	271	325	337	127	127
	SY055	185	185	238	242	258	262	205	208	235	243	325	337	0	0
	SY056	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY057	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY058	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY061	185	187	232	240	258	260	205	208	247	247	0	0	127	127
	SY062	175	175	242	242	258	262	208	208	255	261	337	337	127	147
	SY063	175	175	0	0	0	0	202	205	255	261	325	337	129	129
	SY064	0	0	238	240	256	262	205	208	255	261	325	337	127	127
	SY067	175	175	238	242	258	262	205	208	253	261	325	337	121	127
	SY068	175	175	238	242	258	262	205	208	255	261	325	337	129	129
	SY070	175	175	0	0	0	0	202	205	255	261	325	337	129	129
	SY071	175	175	238	242	258	262	205	208	255	263	325	337	127	127
	SY074	175	175	238	242	258	262	208	208	255	269	325	337	127	127
	SY075	175	175	224	226	0	0	205	208	255	261	325	337	127	127
	SY077	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY079	185	187	238	242	258	258	205	208	267	269	325	337	127	147

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2008 ✂	SY080	175	175	236	242	258	262	205	208	255	261	325	337	129	129
	SY081	175	175	0	0	258	262	205	208	253	259	325	337	127	127
	SY082	175	175	236	242	258	262	205	208	255	261	325	337	127	127
	SY083	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY085	175	175	238	242	258	262	205	208	255	259	325	337	127	127
	SY086	175	187	238	242	258	262	208	208	255	269	325	337	127	127
	SY087	175	175	236	240	258	262	205	208	255	261	325	337	115	127
	SY088	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY089	175	187	260	260	0	0	205	208	255	261	325	337	127	127
	SY090	175	175	238	238	258	262	205	208	255	267	325	337	119	119
	SY091	175	187	236	240	258	262	208	208	255	271	325	337	0	0
	SY095	175	175	0	0	0	0	205	208	255	261	325	337	127	127
	SY097	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY098	175	187	238	242	258	262	208	208	247	271	325	337	127	147
	SY099	175	175	238	240	258	262	205	208	253	261	325	337	129	129
	SY100	177	183	238	238	258	262	205	208	255	267	325	337	119	119
	SY101	175	175	222	224	258	262	205	208	253	261	325	337	117	127
	SY103	175	175	224	224	258	262	205	208	255	261	325	337	127	127
	SY104	175	175	238	242	258	262	208	208	255	261	0	0	127	127
	SY105	175	183	236	242	258	262	205	208	255	261	325	337	129	129
	SY106	175	175	236	242	258	262	205	208	255	261	325	337	0	0
	SY108	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY109	175	175	236	240	258	262	205	208	255	261	325	337	115	127
	SY110	175	187	236	240	258	262	208	208	255	271	325	337	127	127
	SY112	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY113	175	175	240	240	258	262	205	208	255	267	325	337	117	117
	SY114	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY115	175	175	230	240	258	262	205	205	267	271	337	337	127	147
	SY116	175	187	238	238	258	258	205	208	255	261	325	337	127	127
	SY117	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY118	175	175	238	242	258	262	205	208	257	263	325	337	127	127
	SY119	175	185	238	242	258	262	205	208	253	261	325	337	127	127
	SY120	175	175	238	242	258	262	205	208	267	271	325	337	127	127
	SY122	175	175	236	242	258	262	205	208	255	261	325	337	115	127
	SY123	0	0	236	242	258	262	205	208	255	271	0	0	127	147
	SY124	175	175	0	0	0	0	205	208	253	261	325	337	129	129
	SY125	175	175	0	0	0	0	205	208	255	261	325	337	127	127
	SY126	175	175	238	242	0	0	205	208	255	261	325	337	127	127
	SY127	175	187	238	242	258	262	208	208	255	271	325	337	121	127
	SY130	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY131	177	177	236	240	258	262	205	208	255	261	325	337	127	127
	SY132	175	175	224	226	258	262	205	208	255	261	325	337	127	127
	SY133	175	175	236	240	258	262	205	208	255	261	325	337	121	127
	SY134	175	175	236	240	258	262	205	208	253	261	325	337	127	127
	SY136	187	187	242	242	250	258	205	208	267	267	325	337	147	147
	SY137	175	187	236	242	258	262	205	208	255	261	325	337	127	127
	SY138	181	185	238	238	258	258	208	208	255	255	307	337	117	127
	SY140	175	175	0	0	0	0	205	208	253	261	325	337	129	129

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2008 ✕	SY141	175	175	0	0	0	0	199	205	0	0	325	337	117	129
	SY142	175	175	238	242	258	262	205	208	253	261	325	337	127	127
	SY143	0	0	0	0	0	0	208	208	255	267	325	337	127	127
	SY145	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY146	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY147	175	175	236	240	258	262	205	208	255	261	325	337	115	127
	SY152	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY153	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY154	175	175	0	0	0	0	208	208	253	261	325	337	129	129
	SY155	175	187	238	242	258	262	208	208	255	271	325	337	127	127
	SY156	175	175	240	240	258	262	205	208	255	261	325	337	127	127
	SY157	175	175	236	242	258	262	205	208	255	261	325	337	121	127
	SY159	175	175	236	242	258	262	202	205	253	261	325	337	0	0
	SY148	175	175	238	240	258	262	205	208	255	261	325	337	127	127
	SY149	175	175	236	242	258	262	205	208	255	261	325	337	119	127
	SY151	175	175	236	242	256	262	205	208	255	261	325	337	127	127
	SY160	175	175	238	242	258	262	205	208	257	261	325	337	127	127
	SY161	177	177	238	242	258	262	205	208	255	261	325	337	127	127
	SY162	173	173	238	242	258	262	205	208	267	271	325	337	127	127
	SY167	175	177	236	240	258	262	205	208	255	261	325	337	127	127
	SY168	175	175	240	242	258	262	205	208	255	261	325	340	115	127
	SY169	175	183	236	240	258	262	205	208	255	261	325	337	127	127
	SY170	175	175	236	240	258	262	205	208	255	261	325	337	119	127
	SY171	175	187	238	242	250	258	208	208	247	267	325	337	127	147
	SY172	175	175	236	242	258	262	205	208	255	261	325	337	127	127
	SY173	175	175	236	242	258	262	205	208	255	261	325	337	127	127
	SY174	175	175	238	240	258	262	205	208	255	261	325	337	127	129
	SY179	175	187	242	244	258	262	205	208	255	271	307	325	117	127
	SY182	175	175	236	240	256	260	205	208	255	261	325	337	127	127
	SY183	175	175	238	242	258	262	205	208	253	261	325	337	127	127
	SY184	175	187	238	242	258	262	205	208	267	271	325	337	127	127
	SY185	175	175	236	242	258	262	205	208	255	261	325	337	127	127
	SY186	175	175	236	240	258	260	205	208	255	261	325	337	127	127
	SY188	175	175	238	238	258	262	205	208	255	261	325	337	117	127
	SY190	187	187	242	244	258	258	205	205	263	263	325	337	129	143
	SY191	175	175	0	0	0	0	202	205	253	261	325	337	131	131
	SY192	175	175	236	242	258	262	205	208	253	261	325	337	127	127
	SY193	175	175	0	0	0	0	205	208	253	261	325	337	129	129
	SY195	175	175	236	240	258	262	205	205	255	261	325	337	127	127
	SY196	185	187	238	238	258	258	205	208	267	271	325	337	127	147
	SY199	177	183	238	242	258	284	205	208	255	261	325	337	127	127
	SY200	175	175	238	242	292	292	205	208	255	261	325	337	129	129
	SY201	175	175	238	242	258	262	205	208	255	261	325	337	119	127
	SY202	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY203	175	183	238	242	258	262	205	208	253	261	325	337	129	129
	SY206	175	175	236	240	258	262	202	208	255	261	325	337	127	127
	SY207	175	175	238	242	258	258	205	208	255	261	325	337	127	127
	SY208	175	175	238	242	258	262	205	208	255	261	325	337	129	129

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2008 ♂	SY209	175	175	224	224	256	262	205	208	253	261	325	337	127	127
	SY210	0	0	236	242	258	262	202	205	255	261	325	337	0	0
	SY211	175	183	0	0	258	262	205	208	253	259	325	337	129	129
	SY212	175	175	246	250	0	0	205	208	253	261	325	337	127	127
	SY213	175	187	238	242	258	262	205	208	265	269	325	337	127	127
	SY214	175	175	236	242	258	262	205	208	255	261	319	337	127	127
	SY216	175	175	236	240	258	260	205	208	255	261	325	337	127	127
	SY217	175	175	244	244	258	262	205	208	255	261	325	337	127	127
	SY218	175	175	238	242	258	262	205	208	253	261	325	337	127	127
	SY219	175	175	236	240	258	262	205	208	255	261	325	337	125	129
	SY220	173	185	0	0	0	0	205	205	255	271	325	337	127	127
	SY221	175	183	238	240	256	262	205	208	255	261	325	337	127	127
	SY228	175	175	0	0	258	262	205	208	255	261	325	337	0	0
	SY222	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY226	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY227	175	175	238	242	258	262	205	208	255	261	325	337	121	127
	SY231	175	175	236	242	258	262	205	208	255	261	325	337	115	127
	SY234	175	175	236	244	258	262	205	208	255	261	325	337	127	127
	SY237	175	175	238	238	258	262	205	208	255	265	325	337	115	117
	SY238	175	175	238	242	258	262	205	208	255	261	325	337	119	127
	SY239	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY240	175	175	238	242	258	262	205	208	253	261	325	337	119	127
	SY241	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY243	175	187	242	242	258	262	205	208	255	267	325	337	127	127
	SY244	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY246	175	175	236	240	258	262	205	208	255	261	325	337	119	127
	SY247	175	187	0	0	0	0	208	208	255	271	325	337	129	129
	SY248	175	187	238	242	258	262	205	208	255	261	325	337	127	127
	SY249	175	175	238	238	258	262	205	208	255	267	325	337	117	117
	SY250	175	175	238	242	258	262	208	208	255	267	325	337	117	117
	SY251	177	183	238	238	258	262	205	208	255	261	325	337	119	127
	SY252	175	175	238	238	258	262	205	208	255	267	325	337	117	117
	SY253	175	175	238	238	258	262	205	208	255	267	325	337	117	117
	SY254	185	185	238	242	258	262	205	208	255	261	325	337	127	127
	SY255	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY256	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY259	175	183	238	242	258	262	205	208	255	261	325	337	127	127
	SY274	175	175	0	0	0	0	205	208	253	259	325	337	129	129
	SY278	175	187	238	242	258	262	205	208	267	271	325	337	115	127
	SY279	175	175	238	240	258	262	205	208	255	261	325	337	119	127
	SY280	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY281	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY282	175	175	238	238	262	262	208	208	247	255	337	337	119	127
	SY285	175	175	0	0	0	0	205	208	255	261	325	337	129	129
	SY288	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY289	175	175	238	240	258	262	205	208	253	265	325	337	117	117
	SY290	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY292	175	183	236	242	258	262	205	208	255	261	325	337	127	127

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2008 ✂	SY294	0	0	238	238	258	262	208	208	247	265	325	325	117	129
	SY296	0	0	238	242	258	262	208	211	253	269	322	322	129	129
	SY297	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY299	175	175	238	240	256	260	205	208	255	267	325	337	117	117
	SY300	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY301	175	175	236	240	258	262	205	208	255	261	325	337	147	147
	SY302	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY304	175	187	236	236	262	262	208	208	255	255	313	319	115	117
	SY305	175	175	238	242	258	262	205	208	255	261	325	337	119	127
	SY306	175	175	238	242	258	262	205	208	255	261	325	337	127	127
	SY307	175	175	236	240	258	262	205	208	255	261	325	337	127	127
	SY308	175	175	236	240	256	260	205	208	255	261	325	337	127	127
	SY309	175	187	242	242	258	258	205	208	267	271	337	337	127	147
	SY310	175	175	236	240	258	262	205	208	255	261	325	337	127	129
2009 ☼	SY317	175	187	238	238	258	262	208	208	257	257	325	337	117	117
	SY318	175	175	238	238	258	262	208	208	249	255	319	337	117	117
	SY319	175	175	238	242	258	262	208	208	255	255	307	319	119	127
	SY321	175	175	238	242	258	262	208	208	247	255	307	325	117	127
	SY322	175	187	238	240	258	262	208	208	247	267	307	319	117	127
	SY324	175	187	238	238	258	262	208	208	247	255	325	337	117	127
	SY325	175	175	238	238	262	262	208	208	249	255	307	337	117	127
	SY326	175	175	238	238	258	262	205	205	249	249	319	325	119	127
	SY327	175	175	238	238	262	262	208	208	249	267	325	325	117	147
	SY328	175	187	236	238	262	262	208	208	249	249	325	337	117	117
	SY329	175	185	238	242	258	262	205	208	249	255	325	325	117	147
	SY330	175	175	238	238	262	262	208	208	249	255	325	325	117	117
	SY331	175	175	238	238	262	262	208	208	249	255	307	319	117	147
	SY333	175	187	230	242	262	262	208	208	253	265	319	325	127	147
	SY334	175	175	238	238	258	262	208	208	267	267	319	319	169	169
	SY335	175	187	230	238	262	262	208	208	255	255	319	325	119	127
	SY336	175	187	238	240	262	262	208	208	249	267	325	337	117	145
	SY337	175	175	238	238	258	262	208	208	249	255	319	325	127	127
	SY338	175	187	238	240	262	262	208	208	249	267	325	337	115	147
	SY339	175	187	238	238	258	258	208	208	247	267	307	337	115	149
	SY341	175	185	238	238	262	262	208	208	247	255	319	325	117	117
	SY342	175	175	238	238	262	262	205	205	267	267	325	337	119	119
	SY343	175	187	236	238	258	262	205	208	249	255	319	337	115	125
	SY344	175	175	238	238	258	264	208	235	247	247	325	337	145	145
	SY346	175	187	238	240	258	262	208	208	267	267	319	319	115	127
	SY347	175	187	238	238	258	262	208	208	255	265	319	337	115	125
	SY348	175	175	230	238	262	262	208	208	255	255	325	337	127	147
	SY349	175	187	238	238	262	262	208	208	249	255	337	337	117	119
	SY350	175	175	224	238	262	262	208	208	247	255	307	337	119	127
	SY353	175	187	238	242	262	262	208	208	249	255	319	325	117	127
	SY354	175	175	238	238	262	262	208	208	249	255	319	325	117	147
	SY355	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY356	175	175	238	238	258	262	208	208	249	267	325	337	119	147
	SY357	175	187	238	238	262	262	205	205	249	249	325	337	117	125

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2009 ☼	SY358	175	187	238	238	262	262	208	208	249	267	307	337	117	147
	SY359	175	175	224	238	258	258	205	208	255	267	319	325	117	127
	SY360	175	175	238	242	258	262	208	208	267	267	319	325	117	125
	SY361	175	175	230	238	262	262	208	208	255	255	319	325	117	127
	SY362	175	175	238	238	258	262	208	208	249	255	325	325	117	127
	SY364	175	187	238	242	258	262	208	208	249	255	325	337	119	147
	SY365	187	187	238	238	262	262	211	211	255	267	307	319	115	127
	SY366	175	175	236	236	262	262	208	208	249	249	325	325	117	117
	SY367	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY368	175	175	238	242	258	262	208	208	255	255	319	325	115	127
	SY369	175	175	238	238	262	262	208	208	249	255	325	337	127	147
	SY371	175	187	238	238	258	258	208	208	257	273	319	337	115	125
	SY373	175	187	238	238	262	262	208	208	249	255	319	325	115	115
	SY374	175	187	224	238	258	262	208	208	249	267	307	319	117	127
	SY375	175	187	238	240	258	262	208	208	249	267	307	319	115	127
	SY377	175	187	238	238	258	258	0	0	255	271	0	0	117	127
	SY378	175	187	236	238	262	262	208	208	249	249	325	337	117	117
	SY379	175	175	230	238	262	262	208	208	265	265	319	325	117	125
	SY380	175	187	224	238	258	262	208	208	249	255	307	319	117	117
	SY381	175	175	238	238	262	262	208	208	249	255	307	337	117	127
	SY382	173	173	238	238	258	262	208	208	249	267	307	319	117	127
	SY383	175	175	238	238	258	280	208	208	247	255	325	337	117	117
	SY384	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY385	175	175	238	240	258	258	208	208	267	267	319	325	117	127
	SY386	175	187	230	238	258	262	208	208	255	267	307	319	117	127
	SY388	175	175	238	238	262	262	208	208	249	249	325	325	117	117
	SY389	175	175	238	238	258	262	208	208	249	255	319	325	115	127
	SY391	175	187	238	238	258	258	208	208	255	271	319	337	115	127
	SY392	175	187	230	238	262	262	208	208	249	255	325	337	117	147
	SY393	175	175	238	238	262	262	208	208	249	267	325	325	119	127
	SY394	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY395	175	187	238	238	258	258	211	211	255	271	319	337	115	125
	SY396	175	187	238	238	262	262	208	208	249	267	319	319	117	127
	SY397	175	187	242	242	258	262	205	205	259	271	325	337	127	127
	SY398	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY399	175	187	238	238	258	258	208	208	255	271	319	337	115	127
	SY400	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY401	175	187	218	238	262	262	208	208	249	267	319	325	115	125
	SY402	175	187	238	238	258	258	208	208	255	271	319	337	115	127
	SY403	175	187	238	238	258	258	205	205	259	275	319	337	115	127
	SY404	175	187	238	238	258	262	0	0	255	265	0	0	127	127
	SY405	175	187	238	240	262	262	205	205	253	265	319	337	115	125
	SY406	175	187	238	240	262	262	205	205	255	267	319	337	115	127
	SY407	175	187	238	242	262	262	208	208	255	267	319	337	115	127
	SY408	175	187	218	238	258	258	208	208	255	267	319	319	117	127
	SY410	175	187	230	238	258	262	208	208	255	255	319	337	115	127
	SY411	175	175	238	238	258	262	208	208	249	255	0	0	117	127

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2009 ♂	SY412	175	187	240	240	258	262	208	208	255	255	319	337	117	127
	SY413	175	185	240	240	258	262	208	208	255	255	319	337	117	127
	SY415	175	187	238	238	258	262	208	208	255	271	337	337	127	127
	SY416	175	187	238	238	258	262	208	208	249	255	337	337	119	127
	SY418	175	175	238	238	258	262	205	205	247	255	319	325	119	127
	SY419	185	185	238	242	262	262	205	208	255	267	319	337	117	147
	SY420	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY421	175	187	238	242	262	262	205	208	267	271	325	337	119	127
	SY426	175	187	238	238	258	262	208	208	247	247	325	337	117	127
	SY428	175	187	240	240	262	262	208	208	249	249	325	337	117	127
	SY429	175	187	238	240	258	262	208	208	249	271	0	0	117	127
	SY430	175	187	238	238	258	262	208	208	255	267	325	337	117	127
	SY431	175	175	238	242	262	262	208	208	255	267	319	337	119	127
	SY432	175	187	238	238	258	262	208	208	255	267	325	337	117	127
	SY433	175	187	236	240	258	262	205	208	269	271	325	337	127	127
	SY434	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY435	175	187	236	240	258	262	205	205	247	265	325	340	117	127
	SY436	175	187	238	242	258	262	205	205	247	265	325	340	117	127
	SY437	175	187	238	242	258	262	208	208	249	267	325	337	115	127
	SY438	175	187	238	242	258	262	208	208	249	267	325	340	115	123
	SY439	175	187	240	240	260	260	208	208	255	271	319	337	115	127
	SY440	175	185	236	240	258	262	208	208	249	267	325	337	117	127
	SY441	175	187	238	238	258	258	205	205	255	271	319	340	117	127
	SY442	175	187	238	238	258	262	208	208	255	255	319	337	117	127
	SY443	175	187	238	238	258	262	208	208	255	255	319	337	117	127
	SY444	175	187	238	242	262	262	205	205	229	255	325	337	117	127
	SY445	175	187	238	238	256	264	208	208	255	255	319	337	117	127
	SY446	175	187	238	238	258	262	205	205	253	253	319	337	117	127
	SY447	175	187	238	242	258	262	208	208	249	267	325	337	117	127
	SY448	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY449	175	187	238	238	262	262	208	208	255	267	325	337	117	147
	SY450	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY451	175	187	238	242	258	262	208	208	247	267	307	325	117	119
	SY452	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY453	175	187	236	240	258	262	208	208	249	267	325	337	117	127
	SY456	175	175	238	238	258	262	205	205	255	255	325	325	117	127
	SY457	175	187	240	240	258	258	208	208	255	271	319	337	117	127
	SY458	175	187	238	242	258	262	208	208	249	271	325	325	127	127
	SY459	175	187	238	238	258	258	208	208	255	271	319	337	115	127
	SY460	175	187	232	238	262	262	205	205	247	253	319	340	115	127
	SY461	175	175	236	236	262	262	208	208	249	267	319	337	117	127
	SY462	175	187	238	242	258	262	208	208	247	267	0	0	117	127
	SY463	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY464	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY465	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY466	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY467	175	187	238	238	258	262	208	208	255	255	319	337	117	127

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2009 ✂	SY468	175	187	238	238	258	258	208	208	253	271	319	337	115	125
	SY469	175	185	238	238	258	258	205	205	253	269	319	337	117	127
	SY470	175	187	238	238	258	258	208	208	253	269	319	337	117	127
	SY471	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY472	175	187	238	238	258	258	208	208	249	249	319	337	117	127
	SY473	175	187	238	238	258	262	208	208	255	267	319	337	117	127
	SY474	175	187	238	238	258	258	208	208	255	273	319	337	117	127
	SY475	175	187	238	238	258	262	208	208	255	255	319	337	117	127
	SY476	175	187	238	238	258	262	208	208	255	267	325	337	117	127
	SY477	175	187	238	238	262	262	208	208	249	255	319	337	119	127
	SY478	175	187	238	238	262	262	208	208	249	255	319	337	119	127
	SY479	175	175	238	238	258	262	208	208	249	255	319	319	117	147
	SY480	175	187	238	238	262	262	208	208	249	255	325	337	119	127
	SY481	175	187	238	238	258	258	208	208	255	267	319	319	117	119
	SY482	175	187	236	240	258	262	208	208	249	267	325	337	117	127
	SY483	175	175	238	242	258	258	202	205	265	269	325	337	117	127
	SY485	175	187	238	238	258	262	208	208	255	267	325	337	127	147
	SY486	175	187	238	238	258	262	208	208	255	255	319	337	117	127
	SY487	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY488	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY490	175	187	238	238	262	262	208	208	249	249	325	337	117	119
	SY491	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY493	175	175	238	242	0	0	208	208	0	0	319	337	0	0
	SY494	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY496	175	187	236	236	258	258	205	205	0	0	319	337	117	127
	SY499	175	187	238	238	258	258	205	205	253	271	319	337	117	127
	SY501	175	187	238	238	258	258	205	205	253	271	319	337	117	127
	SY504	175	187	238	238	258	258	205	208	255	271	0	0	117	127
	SY505	175	175	230	236	262	262	208	208	249	267	319	325	117	147
	SY508	175	175	230	236	262	262	205	205	249	267	319	325	117	147
	SY510	175	187	238	238	258	258	205	205	255	271	319	337	115	127
	SY512	175	175	238	238	262	262	208	208	0	0	319	325	0	0
	SY513	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY514	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY516	175	187	238	238	258	258	208	208	253	269	319	337	117	127
	SY517	175	187	238	238	258	258	205	205	255	269	319	337	117	127
	SY519	175	187	238	238	258	258	205	205	253	271	319	337	117	127
	SY521	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY522	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY523	175	187	238	242	258	262	208	208	255	267	325	337	117	127
	SY525	175	187	238	238	258	258	205	205	255	271	319	337	117	127
	SY527	175	185	230	236	258	262	205	205	267	275	319	325	117	147
	SY533	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY534	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY537	175	187	238	238	262	262	208	208	267	267	325	325	117	127
	SY538	175	187	238	238	260	264	208	208	251	265	319	340	117	127
	SY539	175	187	238	238	262	262	208	208	267	267	0	0	0	0

Appendix 2. Continued.

	Strain	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
2009 ✂	SY540	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY541	175	187	236	242	258	262	202	208	253	269	307	325	117	127
	SY542	175	187	238	238	258	262	208	208	255	267	319	337	117	127
	SY543	175	175	238	242	258	262	205	208	255	267	319	325	117	127
	SY544	175	187	238	242	258	262	208	208	249	267	325	337	117	127
	SY545	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY547	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY548	175	187	238	238	258	258	208	208	255	255	319	337	117	127
	SY551	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY553	175	175	238	238	258	262	208	208	249	255	325	325	119	127
	SY554	175	175	238	238	262	262	208	208	249	255	325	337	117	119
	SY555	175	187	238	244	258	262	205	208	255	271	0	0	117	127
	SY558	175	175	238	238	262	262	208	208	249	255	325	325	117	147
	SY561	175	175	238	238	262	262	208	208	249	255	325	325	117	147
	SY562	175	187	238	238	262	262	208	208	249	255	337	337	117	127
	SY563	175	187	230	238	262	262	208	208	249	255	325	337	117	119
	SY564	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY566	175	187	238	238	262	262	208	208	249	249	325	337	117	119
	SY567	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY568	175	175	230	236	258	262	208	208	255	267	325	325	127	147
	SY570	175	187	238	242	258	264	208	208	249	255	325	337	119	127
	SY572	175	185	238	242	258	258	208	208	0	0	319	337	117	127
	SY575	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY576	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY577	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY578	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY579	175	187	238	238	262	262	205	208	267	267	319	337	117	127
	SY580	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY581	175	187	238	242	258	262	208	208	249	267	325	337	117	127
	SY582	175	185	236	238	258	258	208	208	255	271	319	337	117	127
	SY583	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY584	175	187	240	240	262	262	208	208	249	255	325	337	117	117
	SY586	175	187	238	238	258	258	208	208	255	271	319	337	117	127
	SY587	175	187	268	242	258	262	208	208	249	267	325	337	117	127
	SY588	175	187	238	238	258	262	208	208	249	267	319	337	119	127
	SY589	175	175	230	238	262	262	208	208	267	267	319	325	117	127
	SY590	175	187	238	238	258	258	208	208	255	271	319	337	117	127

Appendix 3: Microsatellite and ITS-type patterns of the parental and F1 strains sampled the crossing experiments (Chapter 7). The code of parental and F1 strains, (Strains); the ITS-type (ITS); the allelic combination over the seven microsatellites tested (PNm2, PNm5, PNm7, PNm3, PNm6, PNm16 and PNm1). Box colours indicate parental (pink) and F1 strains (Cross_1, blue; _2, yellow; _3, orange; _4, green).A n.a. denotes the ITS-type 'not-analysed'.

strains	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
SY017	A	175	175	238	242	262	262	205	208	261	267	325	337	127	127
SY278	B	175	187	238	242	258	262	205	208	267	271	325	337	115	127
SY138	B	175	187	238	238	258	258	208	208	255	255	307	337	117	127
SY378	A/B	175	187	238	238	262	262	208	208	249	249	325	337	115	117
SY379	A	175	175	230	238	262	262	208	208	267	267	319	325	117	127
CR05A	A/B	175	187	238	238	262	262	205	205	261	271	337	337	127	127
CR05B	A/B	175	187	238	242	262	262	208	208	267	271	325	337	127	127
CR07B	n.a.	175	187	238	238	258	262	205	208	267	271	337	337	127	127
CR08A	n.a.	175	175	238	238	262	262	205	208	261	267	325	325	127	127
CR08B	A	175	187	238	238	258	262	205	208	261	267	325	337	127	127
CR09A	A/B	175	187	238	238	262	262	205	208	267	271	325	337	127	127
CR11A	n.a.	175	175	238	238	258	262	205	205	261	267	325	337	127	127
CR12A	n.a.	175	187	0	0	0	0	205	208	267	267	337	337	127	127
CR12B	A/B	175	187	242	242	258	262	205	205	261	267	325	337	127	127
CR16A	A/B	175	187	242	242	258	262	205	208	267	271	325	337	127	127
CR17A	A/B	175	187	242	242	258	262	205	208	261	271	337	337	127	127
CR17B	n.a.	175	187	242	242	258	262	205	208	261	271	337	337	127	127
CR20A	A/B	175	187	238	242	258	262	205	208	267	267	325	337	127	127
CR21A	A/B	175	187	242	242	258	262	208	208	261	271	325	337	127	127
CR21B	A/B	175	175	238	238	258	262	205	208	261	267	325	325	127	127
CR22B	A/B	175	187	238	242	258	262	205	208	267	271	337	337	127	127
CR24A	A	175	175	238	238	262	262	205	205	261	267	325	337	127	127
CR25	A/B	175	187	238	242	262	262	205	208	261	271	337	337	127	127
CR27A	A/B	175	187	238	238	262	262	205	205	261	271	337	337	127	127
CR27B	n.a.	175	175	238	238	262	262	205	205	267	267	325	325	127	127
CR6A	n.a.	175	187	242	242	258	262	208	208	267	271	325	337	127	127
CRC2	A/B	175	175	242	242	262	262	205	205	261	267	325	337	127	127
CRD1	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127
CRD2	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127
CRF1	A/B	175	175	238	242	258	262	205	208	267	271	325	337	127	127
CRI1	n.a.	175	187	0	0	262	262	205	208	267	267	0	0	127	127
CRI2	A/B	175	175	238	238	258	262	208	208	267	271	325	337	127	127
CS02A	n.a.	175	187	238	242	258	262	205	205	261	267	325	337	127	127
CS09A	A/B	175	175	238	238	258	262	205	205	261	271	325	337	127	127
CS101A	A/B	175	187	238	242	262	262	205	205	261	267	325	337	127	127
CS103B	n.a.	175	187	238	242	258	262	208	208	267	271	337	337	127	127
CS106A	A/B	175	187	238	242	262	262	205	205	261	267	337	337	127	127
CS108B	A/B	175	187	238	238	258	262	205	208	267	267	325	337	127	127
CS110A	n.a.	175	175	242	242	262	262	208	208	267	267	0	0	127	127
CS112A	n.a.	175	187	238	242	262	262	205	205	261	271	337	337	127	127

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS112B	n.a.	175	187	238	238	258	262	205	208	267	267	325	325	127	127
CS114A	n.a.	175	175	238	242	258	262	208	208	267	267	325	325	127	127
CS115A	A/B	175	175	238	242	258	262	205	208	267	267	325	337	127	127
CS116A	n.a.	175	175	238	242	258	262	205	208	261	271	325	337	127	127
CS117A	n.a.	175	187	238	242	258	262	205	208	267	267	337	337	127	127
CS117B	A/B	175	187	238	242	262	262	205	205	267	267	337	337	127	127
CS119B	n.a.	175	187	238	242	258	262	205	208	267	267	325	337	127	127
CS120A	n.a.	175	175	238	242	262	262	208	208	261	271	325	325	127	127
CS120B	n.a.	175	175	242	242	262	262	205	208	261	267	325	337	127	127
CS14A	A/B	175	187	238	242	258	262	208	208	261	267	325	337	127	127
CS15B	n.a.	175	175	238	242	258	262	205	208	267	267	325	325	127	127
CS41B	A/B	175	175	238	242	258	262	208	208	261	267	325	337	127	127
CS55A	n.a.	175	175	238	242	262	262	205	208	261	267	325	325	127	127
CS58A	A/B	175	175	238	238	258	262	205	208	267	267	325	337	127	127
CS61B	A/B	175	187	238	242	258	262	205	208	267	271	325	325	127	127
CS62A	A/B	175	187	238	242	262	262	205	208	267	271	325	325	127	127
CS62B	A/B	175	175	238	242	258	262	205	208	261	271	325	337	127	127
CS79B	n.a.	175	175	238	238	262	262	205	208	261	271	325	337	127	127
CS80A	A/B	175	175	238	242	258	262	205	208	267	267	325	337	127	127
CS81A	n.a.	175	187	238	238	258	262	205	205	267	271	337	337	127	127
CS81B	A	175	187	238	242	262	262	205	208	267	267	337	337	127	127
CS83A	n.a.	175	187	238	238	258	262	208	208	267	267	337	337	127	127
CS83B	n.a.	175	175	238	242	262	262	208	208	0	0	0	0	127	127
CS85B	n.a.	175	175	242	242	258	262	205	208	261	267	325	337	127	127
CS87A	A/B	175	187	242	242	258	262	205	208	267	271	325	337	127	127
CS89B	n.a.	175	187	238	242	262	262	208	208	267	267	337	337	127	127
CS98A	n.a.	175	187	238	238	262	262	208	208	267	267	337	337	127	127
CS122A	n.a.	175	175	238	238	258	262	208	208	255	267	307	325	119	127
CS125B	n.a.	175	187	238	238	258	262	208	208	255	261	325	337	117	117
CS127B	n.a.	175	187	238	242	258	262	205	208	255	267	337	337	127	127
CS127D	n.a.	175	175	238	242	258	262	208	208	255	267	307	325	119	127
CS128A	A/B	175	187	238	238	258	262	205	208	255	261	307	325	127	127
CS128B	A/B	175	187	238	242	258	262	208	208	255	267	337	337	127	127
CS128C	n.a.	175	187	238	238	258	262	208	208	255	267	0	0	127	127
CS129D	n.a.	175	187	236	238	258	262	205	208	255	261	325	337	127	127
CS130A	n.a.	175	187	238	238	258	262	208	208	255	261	325	337	117	127
CS130C	n.a.	175	187	238	238	258	262	208	208	255	261	325	337	119	127
CS131B	n.a.	175	175	238	242	258	262	208	208	255	261	307	337	127	127
CS131C	n.a.	175	175	238	242	258	262	208	208	0	0	307	337	127	127
CS132B	n.a.	175	175	238	238	258	262	205	208	255	267	337	337	127	127
CS132D	A/B	175	187	238	238	258	262	205	208	255	261	325	337	117	127

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS133A	n.a.	175	175	238	242	258	262	205	208	255	261	325	337	117	127
CS133B	A/B	175	175	238	238	258	262	208	208	255	267	325	337	117	127
CS133C	n.a.	175	187	238	242	258	262	208	208	255	261	325	337	117	127
CS134A	n.a.	175	175	238	240	258	262	208	208	255	267	307	325	117	127
CS134B	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	127	127
CS134C	A/B	175	187	238	238	258	262	205	208	255	267	325	337	117	127
CS134D	A/B	175	175	238	248	258	262	205	208	255	267	307	337	127	127
CS135A	n.a.	175	187	238	242	258	262	208	208	255	267	337	337	127	127
CS135D	n.a.	175	175	238	242	258	262	205	208	255	267	307	325	127	127
CS136A	A/B	175	187	238	238	258	262	205	208	255	267	325	337	127	127
CS136D	n.a.	175	187	238	242	258	262	208	208	255	267	325	337	117	127
CS137C	n.a.	175	175	238	238	258	262	205	208	255	267	307	325	127	127
CS137D	A/B	175	175	238	238	258	262	208	208	255	261	307	325	127	127
CS138A	A/B	175	187	238	242	258	262	205	208	255	261	307	337	127	127
CS138B	A/B	175	175	238	238	258	262	208	208	255	267	325	337	117	127
CS139A	n.a.	175	175	238	242	258	262	208	208	255	261	307	325	127	127
CS139D	n.a.	175	175	238	238	0	0	205	208	255	267	307	325	127	127
CS140C	A/B	175	175	238	242	258	262	205	208	255	267	307	325	127	127
CS141A	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	127	127
CS141B	n.a.	175	175	238	242	258	262	208	208	255	261	307	325	117	127
CS142A	n.a.	175	187	238	242	258	262	205	208	255	267	337	337	127	127
CS142B	n.a.	175	175	238	238	258	262	205	208	0	0	337	337	117	127
CS142D	n.a.	175	187	238	238	258	262	205	208	255	267	337	337	117	127
CS143A	n.a.	175	187	238	238	258	262	205	208	255	261	337	337	117	127
CS143B	n.a.	175	175	238	238	258	262	205	208	255	261	307	337	127	127
CS143C	n.a.	175	175	238	242	258	262	208	208	255	267	307	325	117	127
CS144A	n.a.	175	175	238	242	258	262	208	208	255	267	307	325	117	127
CS144C	n.a.	175	187	238	242	258	262	205	208	255	261	337	337	117	127
CS144D	n.a.	175	175	238	238	258	262	205	208	255	267	307	337	117	127
CS145A	A/B	175	175	238	242	258	262	208	208	255	267	307	325	127	127
CS145B	A/B	175	187	238	242	258	262	205	208	255	267	337	337	117	127
CS145C	n.a.	175	175	0	0	0	0	208	208	255	261	307	337	127	127
CS145D	n.a.	175	187	238	242	258	262	205	208	0	0	337	337	117	127
CS146A	A/B	175	175	238	238	258	262	205	208	255	261	307	337	127	127
CS146B	n.a.	175	187	238	238	258	262	208	208	255	261	325	337	117	127
CS146D	n.a.	175	175	238	242	258	262	208	208	255	267	307	325	127	127
CS147C	A/B	175	175	238	238	258	262	208	208	255	267	337	337	117	127
CS148B	A/B	175	175	238	242	258	262	208	208	255	267	307	337	127	127
CS148D	A/B	175	187	238	238	258	262	208	208	255	267	325	337	117	127
CS149C	n.a.	175	187	238	238	258	262	208	208	0	0	325	337	127	127
CS149D	n.a.	175	175	238	238	0	0	205	208	255	267	307	337	117	127

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS150A	n.a.	175	175	238	242	258	262	205	208	255	261	307	337	117	127
CS150B	n.a.	175	187	238	238	258	262	208	208	255	261	337	337	117	127
CS150D	n.a.	175	175	238	242	258	262	205	208	255	261	307	337	117	127
CS151B	n.a.	175	187	238	242	258	262	208	208	255	267	337	337	117	127
CS152A	A/B	175	187	238	242	258	262	208	208	255	261	337	337	127	127
CS152D	n.a.	175	187	238	238	258	262	208	208	255	261	337	337	127	127
CS153B	n.a.	175	175	238	238	258	262	208	208	255	261	307	325	117	127
CS153C	n.a.	175	187	238	238	258	262	208	208	255	267	337	337	117	127
CS154A	n.a.	175	175	0	0	0	0	208	208	255	261	307	337	127	127
CS154B	n.a.	175	175	238	238	258	262	205	208	255	267	307	325	117	127
CS155B	n.a.	175	187	238	238	258	262	205	208	255	261	337	337	127	127
CS156B	A/B	175	187	238	242	258	262	205	208	255	261	325	337	117	127
CS157D	A/B	175	175	238	238	258	262	208	208	255	261	307	337	127	127
CS158B	n.a.	175	175	238	238	258	262	208	208	255	261	307	325	127	127
CS158C	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	117	127
CS158D	n.a.	175	175	238	242	258	262	205	208	255	261	307	337	127	127
CS159A	n.a.	175	187	0	0	0	0	208	208	0	0	0	0	127	127
CS159B	n.a.	175	187	238	238	258	262	208	208	255	261	325	337	117	127
CS159C	n.a.	175	175	238	238	258	262	205	208	255	267	337	337	127	127
CS160C	A/B	175	175	238	238	258	262	208	208	255	267	307	325	117	127
CS161C	A/B	175	187	238	242	258	262	205	208	255	261	325	337	117	127
CS162A	A/B	175	175	240	240	258	262	208	208	255	261	307	337	127	127
CS162B	A/B	175	175	238	238	258	262	208	208	255	267	307	325	117	127
CS162D	A/B	175	187	238	242	258	262	208	208	255	267	307	337	127	127
CS163B	n.a.	175	175	238	238	258	262	208	208	255	261	307	325	127	127
CS163C	A/B	175	175	238	238	258	262	205	208	255	261	307	337	117	127
CS164D	n.a.	175	187	238	238	258	262	205	208	255	267	307	337	127	127
CS166D	n.a.	175	175	238	238	258	262	205	208	255	267	307	325	117	127
CS167B	A/B	175	187	238	238	258	262	205	208	255	267	337	337	117	127
CS167C	A/B	175	187	238	242	258	262	205	208	255	267	307	325	117	127
CS168B	n.a.	175	175	238	242	258	262	208	208	255	261	325	337	117	127
CS169B	A/B	175	175	238	242	258	262	205	208	255	267	307	337	127	127
CS169D	A/B	175	175	238	242	258	262	205	208	255	267	307	337	127	127
CS170D	n.a.	175	187	238	238	258	262	205	208	255	267	325	337	127	127
CS176D	A/B	175	187	238	242	258	262	205	208	255	267	307	325	127	127
CS179A	n.a.	175	175	238	242	258	262	208	208	0	0	325	337	117	127
CS180A	n.a.	175	175	238	238	258	262	205	208	255	267	307	337	117	127
CS185D	n.a.	175	187	238	238	258	262	208	208	249	255	337	337	119	119
CS188D	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	117	119
CS189A	n.a.	175	175	238	238	258	262	208	208	249	255	307	325	119	119
CS189B	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	117	127

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS192B	A/B	175	187	238	238	258	262	208	208	249	255	307	337	117	127
CS192C	B	175	187	238	238	258	262	208	208	249	255	307	325	117	127
CS195C	A/B	175	187	238	238	258	262	208	208	249	255	325	337	115	127
CS195D	n.a.	175	187	238	238	258	262	208	208	249	255	307	337	119	119
CS196C	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	117	127
CS197E	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	115	127
CS198A	A/B	175	175	238	238	258	262	208	208	249	255	307	325	117	117
CS198B	A/B	175	187	238	238	258	262	208	208	249	255	307	337	117	117
CS201C	n.a.	175	187	238	238	258	262	208	208	247	255	337	337	117	127
CS202C	n.a.	187	187	238	238	258	262	208	208	249	255	325	337	115	127
CS204A	A/B	175	187	238	238	258	262	208	208	249	255	325	337	115	117
CS205C	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	117	127
CS205D	A/B	187	187	238	238	258	262	208	208	249	255	337	337	115	127
CS207C	n.a.	175	187	238	238	258	262	208	208	249	255	337	337	117	117
CS208C	n.a.	175	175	238	238	258	262	208	208	249	255	307	325	115	127
CS214A	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	117	127
CS225C	B	175	187	238	238	258	262	208	208	249	255	325	337	117	127
CS225D	n.a.	175	187	238	238	258	262	208	208	249	255	325	337	115	127
CS226D	A/B	187	187	238	238	258	262	208	208	249	255	307	337	117	117
CS231C	A/B	175	175	238	238	258	262	208	208	249	255	307	325	119	119
CS233C	A/B	175	187	238	238	258	262	208	208	249	255	325	337	117	127
CS238B	n.a.	175	175	238	238	258	262	208	208	249	255	0	0	117	127
CS241A	A/B	175	187	238	238	258	262	208	208	255	267	319	337	127	127
CS241B	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS241C	n.a.	175	187	230	238	258	262	208	208	255	267	325	337	127	127
CS241D	A/B	175	187	230	238	258	262	208	208	255	267	319	337	117	127
CS244C	n.a.	175	187	230	238	258	262	208	208	0	0	325	337	117	117
CS248A	A/B	175	187	238	238	258	262	208	208	255	267	307	319	127	127
CS248C	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	117
CS250B	A/B	175	175	238	238	258	262	208	208	255	267	307	319	117	127
CS251C	A/B	175	187	238	238	258	262	208	208	255	267	319	337	117	127
CS252C	A/B	175	187	238	238	258	262	208	208	255	267	325	337	127	127
CS259C	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	117	117
CS259D	A/B	175	187	238	238	258	262	208	208	255	267	325	337	117	127
CS263B	A/B	175	187	238	238	258	262	208	208	255	267	325	337	127	127
CS264A	n.a.	175	175	238	238	0	0	208	208	255	267	325	337	127	127
CS266D	n.a.	175	175	238	238	258	262	208	208	255	267	307	325	117	117
CS268D	n.a.	175	187	230	238	258	262	208	208	255	267	319	337	127	127
CS271B	A/B	175	187	238	238	258	262	208	208	255	267	319	337	127	127
CS271C	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	117	117
CS271C2	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	117

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS272A	n.a.	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS272A2	A/B	175	175	230	238	262	262	208	208	267	267	319	325	117	127
CS272B	A/B	175	187	230	238	258	262	208	208	255	267	0	0	117	127
CS273C	A/B	175	175	230	238	258	262	208	208	255	267	307	325	127	127
CS273D	n.a.	175	175	230	238	258	262	208	208	255	267	307	319	117	127
CS274A	A/B	175	175	238	238	258	262	208	208	255	267	307	319	117	127
CS274C2	A/B	175	175	230	238	258	262	208	208	255	267	307	325	117	127
CS274D	A/B	175	187	230	238	258	262	208	208	255	267	319	337	117	127
CS275A	A/B	175	175	230	238	258	262	208	208	255	267	307	325	117	127
CS275D	A/B	175	187	238	238	258	262	208	208	255	267	307	325	117	117
CS276A	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS276B	A/B	175	175	230	238	258	262	208	208	255	267	307	319	117	127
CS276C	A/B	175	175	230	238	258	262	208	208	255	267	307	319	117	127
CS276D	n.a.	175	175	238	238	258	262	208	208	255	267	325	337	117	117
CS277A	A/B	175	187	230	238	258	262	208	208	255	267	325	337	127	127
CS277B	A/B	175	175	238	238	258	262	208	208	255	267	307	319	117	127
CS277B2	A/B	175	187	230	238	258	262	208	208	255	267	319	337	117	127
CS278A	A/B	175	187	238	238	258	262	208	208	255	267	325	337	117	117
CS278B2	A/B	175	187	238	238	258	262	208	208	255	267	319	337	117	127
CS278D	A/B	175	175	230	238	258	262	208	208	255	267	307	325	117	117
CS278D2	n.a.	175	187	238	238	258	262	208	208	255	267	319	337	117	127
CS280A	A/B	175	175	230	238	258	262	208	208	255	267	307	319	117	127
CS280B	A/B	175	187	238	238	258	262	208	208	255	267	319	337	117	127
CS280C	A/B	175	187	230	238	258	262	208	208	255	267	319	337	117	127
CS280C2	A/B	175	175	238	238	258	262	208	208	255	267	307	325	117	117
CS280D	A/B	175	175	230	238	258	262	208	208	255	267	307	319	117	127
CS280D2	A/B	175	187	238	238	258	262	208	208	255	267	307	325	117	117
CS281B	A/B	175	175	230	238	258	262	208	208	255	267	307	325	117	117
CS281D	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS282C	n.a.	175	187	238	238	258	262	208	208	255	267	319	337	117	127
CS282C2	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	117
CS283B	A/B	175	187	238	238	258	262	208	208	255	267	325	337	127	127
CS283D	A/B	175	187	238	238	258	262	208	208	255	267	325	337	117	127
CS284A	n.a.	175	175	230	238	258	262	208	208	255	267	307	325	117	117
CS284A2	A/B	175	175	230	238	258	262	208	208	255	267	307	319	117	117
CS284D	A/B	175	187	230	238	258	262	208	208	255	267	325	337	127	127
CS284D2	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS286A	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS286D	n.a.	175	175	230	238	258	262	208	208	255	267	307	319	117	117
CS287C	A/B	175	187	230	238	258	262	208	208	255	267	325	337	117	127
CS289C	n.a.	175	187	238	238	258	262	208	208	255	267	325	337	117	127

Appendix 3: Continued.

Strain	ITS	*PNm2*		*PNm5*		*PNm7*		*PNm3*		*PNm6*		*PNm16*		*PNm1*	
CS290A	n.a.	175	175	238	238	258	262	208	208	255	267	307	319	117	127
CS292D	A/B	175	175	238	238	258	262	208	208	255	267	307	319	117	117
CS294C	n.a.	175	187	230	238	258	262	208	208	255	267	325	337	117	117
CS294D	n.a.	175	175	238	238	258	262	208	208	255	267	307	319	117	117
CS295A	n.a.	175	175	230	238	258	262	208	208	255	267	307	319	117	117
CS295B	n.a.	175	187	230	238	258	262	208	208	0	0	307	319	0	0

Appendix 4: Statistical comparison of ITS-proportion per ITS-type (-A, -B and -A/B) and over early-summer (☼) and autumnal (✂) blooms 2004 to 2009. Statistical method: comparison of proportion with a threshold error of alpha equal 5%. Grey box indicates no statistical viable data; Blue box indicates that null hypothesis is valid (Ho: there is no significant difference between sample periods) and Yellow box indicate the validity of alternative hypothesis (H1: there is a significant difference between sample periods).

α = 5%		**ITS-A/B type**					
Years		2004 ✂	2005 ✂	2006 ✂	2008	2009☼	2009 ✂
	2004 ✂						
	2005 ✂						
	2006 ✂						
	2008					**H1**	**H1**
	2009 ☼						**Ho**
	2009 ✂						
α = 5%		**ITS-B type**					
Years		2004 ✂	2005 ✂	2006 ✂	2008	2009☼	2009 ✂
ITS-A type	2004 ✂						
	2005 ✂						
	2006 ✂						
	2008					**H1**	**H1**
	2009 ☼						**Ho**
	2009 ✂					**Ho**	

Appendix 5: Statistical comparison of ITS-A/B proportion per subclasses (A-dominant, B-dominant and A/B-mix) between 2008 and 2009 blooms. Statistical method: comparison of proportion with a threshold error of alpha equal 5%. Yellow boxes indicate the validity of alternative hypothesis (H1: there is a significant difference between sample periods).

A/B type	A-dominant	B-dominant	AB-mix
Year1	2008	2008	2008
Year2	2009	2009	2009
n1	12	126	7
n2	36	6	12
N1	167	167	167
N2	98	98	98
p1	0.072	0.754	0.042
p2	0.367	0.061	0.122
pc	0.181	0.498	0.072
Zc	6.030	10.896	2.453
Ualf/2	1.96	1.96	1.96
if Zc > U (H1)	H1	H1	H1

Appendix 6: Statistical comparison of ITS-proportion per ITS-type (-A, -B and -A/B) and over early-summer (☼) and autumnal (🍂) blooms 2004 to 2009. Statistical method: comparison of proportion per ITS-type between sampling period (Year1 and Year2) with a threshold error of alpha equal 5% (U α/2 = 1.96). The number of strains per ITS-type and sampling period (n), the total number of strains per sampling period (N), the proportion of n strains in N (p), the statistical common proportion (pc) and the value for Z-test (Zc). The null hypothesis (Ho) can not be rejected under α=5% if Zc inferior to Uα/2 (blue box), otherwise H1 is accepted (yellow box). Grey box indicates no statistical viable values (N<30 and Np<5).

ITS-A-type							**ITS-B-type**					
Year1	04🍂	05🍂	06🍂	08	08	09 ☼	04🍂	05🍂	06🍂	08	08	09 ☼
Year2	05🍂	06🍂	08	09 ☼	09🍂	09🍂	05🍂	06🍂	08	09 ☼	09🍂	09🍂
n1	3	4	1	1	2	9	2	3	1	21	20	15
n2	4	1	1	9	5	5	3	1	21	15	15	15
N1	6	16	11	167	167	54	6	16	11	167	167	54
N2	16	11	167	54	44	44	16	11	167	54	44	44
p1	0.5	0.25	0.09	0.01	0.01	0.17	0.33	0.19	0.09	0.13	0.12	0.28
p2	0.25	0.09	0.01	0.17	0.11	0.11	0.19	0.09	0.13	0.28	0.34	0.34
N1p1	3	4	1	1	2	9	2	3	1	21	20	15
N2p2	4	1	1	9	5	5	3	1	21	15	15	15
pc	0.32	0.19	0.01	0.05	0.03	0.14	0.23	0.15	0.12	0.16	0.17	0.31
Zc	1.12	1.05	2.59	4.94	3.35	0.75	0.73	0.69	0.34	2.63	3.51	0.67

ITS-A/B-type											
Year1	04🍂	05🍂	05🍂	05🍂	05🍂	06🍂	06🍂	06🍂	08	08	09 ☼
Year2	05🍂	06🍂	08	09 ☼	09🍂	08	09 ☼	09🍂	09 ☼	09🍂	09🍂
n1	1	9	9	9	9	9	9	9	145	145	30
n2	9	9	145	29	24	145	30	24	29	24	24
N1	6	16	16	16	16	11	11	11	167	167	54
N2	16	11	167	54	44	167	54	44	54	44	44
p1	0.17	0.56	0.56	0.56	0.56	0.82	0.82	0.82	0.87	0.87	0.56
p2	0.56	0.82	0.87	0.54	0.55	0.87	0.56	0.55	0.54	0.55	0.55
N1p1	1	9	9	9	9	9	9	9	145	145	30
N2p2	9	9	145	29	24	145	30	24	29	24	24
pc	0.46	0.67	0.84	0.54	0.55	0.87	0.60	0.60	0.79	0.80	0.55
Zc	1.66	1.39	3.20	0.18	0.12	0.47	1.62	1.65	5.17	4.77	0.10

Appendix 7: Statistical comparison of ITS-proportion per ITS-A/B-types (A-dominant versus A-dominant) and over early-summer (☼) and autumnal (✂) blooms 2004 to 2009. Statistical method as explained in Appendix 6.

A/B	A-dominant									
Year1	05 ✂	05 ✂	05 ✂	05 ✂	06 ✂	06 ✂	06 ✂	08	08	09 ☼
Year2	06 ✂	08	09 ☼	09 ✂	08	09 ☼	09 ✂	09 ☼	09 ✂	09 ✂
n1	5	5	5	5	8	8	8	12	12	20
n2	8	12	20	16	12	20	16	20	16	16
N1	9	9	9	9	9	9	9	145	145	29
N2	9	145	29	24	145	29	24	29	24	24
p1	0.56	0.56	0.56	0.56	0.89	0.89	0.89	0.08	0.08	0.69
p2	0.89	0.08	0.69	0.67	0.08	0.69	0.67	0.69	0.67	0.67
pc	0.72	0.11	0.66	0.64	0.13	0.74	0.73	0.18	0.17	0.68
Zc	1.58	4.39	0.74	0.59	6.98	1.19	1.28	7.70	7.13	0.18

Appendix 8: Statistical comparison of ITS-proportion per ITS-A/B-types (A-dominant versus B-dominant) for the autumnal bloom 2008 (unique data for which the test is statistically valid N>30 and Np>5). Statistical method as explained in Appendix 6.

Year	2008
A/B type	**A**-dominant
A/B type	**B**-dominant
n1	12
n2	126
N1	145
N2	145
p1	0.08
p2	0.87
pc	0.48
Zc	13.40

*Appendix 9: Parental strains (SY017, SY138, SY278, SY378 and SY379) frustule size and date record (S_i: frustule size from isolation, S_f: frustule size at each crossing experiment). The size reduction has been monitored and the regression calculates. The * indicates the estimated size value ($S_f = S_i - n_{month} . C_{regression}$) using a mean value of $C_{regression}$ of 0,794 µm.month^{-1}.*

Strain	Size (S_i), net date	Size (S_f), cross date	Size reduction over months (µm ; n_{months})	$C_{regression}$
SY 017	48 µm, Jul08	38 µm, Jul 09 34 µm, Oct 09	(10 ; 12) (14 ; 17)	0.833 0.824
SY 138	36 µm, Sep08	25 µm, Oct 09 25 µm, Nov 09	(11 ; 13) (11 ; 14)	0.846 0.786
SY 278	40 µm, Nov08	36 µm, Jul 09 32 µm, Oct09	(6 ; 8) (8 ; 11)	0.750 0.727
SY 378	48 µm, Jul09	45 µm*, Nov 09	(3*; 4)	-
SY 379	49.8 µm, Jul09	46.8 µm*, Nov 09	(3*; 4)	-

Appendix 10: Groups of multi-F1-strains of microsatellite unique-genotype grouped as a function of their ITS-type. N indicates the number of group per ITS-category (A/B or NR (not-read)) and thereafter groups are separated by a slash bar. Underline strains indicate sister-cells and blue writing indicate a parental strain.

Cross	ITS	N	Groups of same microsatellite genotype pattern
1	A/B	4	CR05A, CR27A / CR16A, CS87A / CRD1, CRD2, CRF1 / CS115A, CS80A
2	A/B	4	CS133B, CS138B / CS156B, CS161C / CS160C, CS162B / CS169B, CS169D
	NR	9	CS127B, CS142A / CS130A, CS146B, CS159B / CS132B, CS159C / CS134B, CS141A / CS143C, CS144A / CS144D, CS180A / CS150A, CS150D / CS154B, CS166D / CS158B, CS163B
3	NR	2	CS189B, CS196C, CS205C, CS214A / CS197E,CS225D
4	A/B	14	SY379, CS272A2 / CS241A, CS271B / CS241B, CS276A, CS281D, CS284D2, CS286A, CS287C / CS241D, CS274D, CS277B2, CS280C / CS248C, CS271C2, CS282C2 / CS250B, CS274A, CS277B / CS251C, CS278B2, CS280B / CS252C, CS263B, CS283B / CS259D, CS283D / CS274C2, CS275A / CS275D, CS280D2 / CS276B, CS276C, CS280A, CS280D / CS277A, CS284D / CS278D, CS281B
	NR	3	CS259C, CS271C / CS278D2, CS282C / CS286D, CS295A

Appendix 11: Mendelian inheritance law test: Cross_1 and Cross_2. Statistical approaches with a comparision of proportion test (N1 and N2: number of total strains consider per locus; n1: number of strain having specific genotype, n2: number of expected strains (n2=N2/2), proportion of genotype (p), pooled proportion (pc), (Zc) value of Z-test, n>30, Np>5, α=5%, Ho null hypothesis accepted, H1 alternative hypothesis accepted if Zc > $U_{\alpha/2}$).

Cross	Cross_1 (SY017 x SY278)						Cross_2 (SY017 x SY138)					
Locus	*PNm2*		*PNm7*		*PNm1*		*PNm2*		*PNm3*		*PNm6*	
Alleles	175 175	175 187	262 262	258 262	115 127	127 127	175 175	175 187	205 208	208 208	255 261	255 267
n1	26	36	24	37	0	62	49	43	43	49	37	49
n2	31	31	30.5	30.5	31	31	46	46	46	46	43	43
N1	62	62	61	61	62	62	92	92	92	92	86	86
N2	62	62	61	61	62	62	92	92	92	92	86	86
p1	0.419	0.581	0.393	0.607	0.000	1.000	0.533	0.467	0.467	0.533	0.430	0.570
p2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500
pc	0.460	0.540	0.447	0.553	0.250	0.750	0.516	0.484	0.484	0.516	0.465	0.535
Zc	0.901	0.901	1.184	1.184	6.429	6.429	0.443	0.443	0.443	0.443	0.917	0.917
$U_{\alpha/2}$	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96
Test	Ho	Ho	Ho	Ho	H1	H1	Ho	Ho	Ho	Ho	Ho	Ho
N1p1	26	36	24	37	0	62	49	43	43	49	37	49
N2p2	31	31	30.5	30.5	31	31	46	46	46	46	43	43

Appendix 12: Mendelian inheritance law test: Cross_3 and Cross_4. Statistical approaches (see Appendix 11).

Cross	Cross_3 (SY378 x SY138)				Cross_4 (SY379 x SY138)					
Locus	*PNm5*		*PNm1*		*PNm2*		*PNm5*		*PNm7*	
Alleles	238 238	238 242	117 127	127 127	175 175	175 187	230 238	238 238	258 262	262 262
n1	49	37	43	45	25	42	37	30	65	1
n2	44.5	44.5	46	46	33.5	33.5	33.5	33.5	66	0
N1	89	89	92	92	67	67	67	67	66	66
N2	89	89	92	92	67	67	67	67	66	66
p1	0.551	0.416	0.467	0.489	0.373	0.627	0.552	0.448	0.985	0.015
p2	0.500	0.500	0.500	0.500	0.500	0.500	0.500	0.500	1.000	0.000
pc	0.525	0.458	0.484	0.495	0.437	0.563	0.526	0.474	0.992	0.008
Zc	0.675	1.128	0.443	0.147	1.481	1.481	0.606	0.606	1.004	1.004
$U_{\alpha/2}$	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96	1.96
Test	Ho	Ho	Ho	Ho	Ho	Ho	Ho	Ho	Ho	Ho
N1p1	49	37	43	45	25	42	37	30	65	1
N2p2	44.5	44.5	46	46	33.5	33.5	33.5	33.5	66	0

Appendix 13: Mendelian inheritance law test _ Statistical approaches with Chi² test (ddl=3 and 4) under alpha 5%: Ho null hypothesis accepted, H1 alternative hypothesis accepted, Δ rare event).

	Cross_1				Cross_2	Cross_3			Cross_4	
PNm	*5*	*3*	*6*	*16*	*16*	*2*	*16*	*1*	*16*	*1*
Chi2	2.13	0.58	1.62	2.34	0.49	4.15	7.16	7.54	8.79	1.93
α = 5%	5.99	5.99	7.81	5.99	7.81	7.81	7.81	7.81	7.81	5.99
Test	Ho	Ho	Ho	Ho	Ho	Ho	Ho	Ho	H1	Ho

*Appendix 14: Microsatellite genotypes record in function of SY017 and SY138, the cross involved (Cross_1 black, Cross_2 green, Cross_3 blue, Cross_4 orange) and the ITS-type (*AB, **B ITS-type belonging) (GIMLET). Underlined strains are sister-cells; Italic strains are neighbour strain (i.e. isolated from the same well); bold-writing strain is a parental strain. Each group is separated by a slash bar.*

	microsatellite similar genotype groups
Female Crosses _1 & _2 (ITS-A x ITS-B types) - Total of 27 groups - Cross_1: 7 groups (black) - Sister-cells underlined. - Cross_2: 20 groups (green)	CR05A*, CR27A*/ CR16A*, CS87A*/ CR17A, CR17B/ CR20A, CS119B/ CRD1*, CRD2*, CRF1*/ CS115A*, CS80A*/ CS62B, CS116A/ CS128B, CS135A/ CS133B*, CS138B*/ CS136A, CS170D/ CS137D, CS158B, CS163B / CS140C, CS135D/ CS145A, CS146D/ CS146A, CS143B/ CS148D, CS158C/ CS156B*,CS161C*/ CS160C*, CS162B*/ CS167B, CS142D/ *CS169B*, CS169D**/ CS127B, CS142A / CS130A, CS146B, CS159B / CS132B, CS159C/ CS134B, CS141A/ CS143C, CS144A/ CS144D, CS180A/ CS150A, CS150D/ CS154B,CS166D
Male SY138 crosses groups (Crosses _2, _3 and _4) Crosses_2 & _4 (ITS-A x ITS-B) Cross_3 (ITS-A x ITS-AB) - Total of 39 groups - Mixed: 2 groups (underlined) and 1 group with female - Cross_2: 18 groups - Cross_3: 3 groups - Cross_4: 15 groups	CS128B, CS135A/ CS133B*, CS138B*/ CS136A, CS170D/ CS137D, CS158B, CS163B/ CS140C, CS135D/ CS145A, CS146D/ CS146A,CS143B/ CS156B*, CS161C*/ CS160C*, CS162B*/ CS167B, CS142D/ CS169B*, CS169D*/ CS127B,CS142A/ CS130A, CS146B,CS159B / CS132B, CS159C CS143C, CS144A/ CS144D, CS180A/ CS150A, CS150D/ CS154B, CS166D/ CS148D*, CS158C, CS259D*, CS283D*, CS289C / CS134B, CS141A, CS252C*, CS263B*, CS283B* / CS195C, CS197E, CS225D/ CS231C, CS189A/ CS233C, CS225C, CS189B, CS196C, CS205C,CS214A/ **SY379***, CS272A2*/ CS241A*, CS271B*/ CS241B*, CS276A*, CS281D*, CS284D2*, CS286A*, CS287C*, CS272A/ CS241D*, CS274D*, CS277B2*, CS280C*/ CS248C*, CS271C2*, CS282C2*, CS294C/ CS250B*, CS274A*, CS277B*, CS290A/ CS251C*, CS278B2*, CS280B*, CS278D2, CS282C/ CS274C2*, CS275A* / CS275D*, CS280D2*/ CS276B*, CS276C*, CS280A*, CS280D*, CS273D/ CS277A*, CS284D*, CS241C/ CS278A, CS259C, CS271C/ CS278D*, CS281B*, CS284A/ CS280C2, CS266D/ CS284A2,CS286D, CS295A/ CS292D, CS294D

*Appendix 15: Genetic structure of the F1 strains from the four crosses (STRUCTURE), grouped in function of their contribution (>50%) to one estimated subpopulation structure. The * and ** indicate the strain whose genotype gather in GIMLET (Appendix 14). Parameter set is: N=231, no missing value, F1 strains only, length of burn-in 10000, MCMC repeats 50000, Admixture models and correlated allele frequencies.*

Group A (50.9-98.3% of contribution)	Intermediary Group
Cross_1: CR05A CR05B CR07B CR08A CR08B CR09A CR11A CR12B CR16A CR17A CR17B CR20A CR21A CR21B CR22B CR24A CR25 CR27A CR27B CR6A CRC2 CRD1 CRD2 CRF1 CRI2 CS02A CS09A CS14A CS159C CS15B CS170D CS41B CS55A CS58A CS61B CS62A CS62B CS79B CS80A CS81A CS81B CS83A CS85B CS87A CS89B CS98A CS101A CS103B CS106A CS108B CS112A CS112B CS114A CS115A CS116A CS117A CS117B CS119B CS120A CS120B Cross_2: CS127B CS132B CS136A CS142A	Cross_2: CS128B CS134B** CS135A CS141A** Cross_4: CS252C** CS263B** CS283B** **&** member of the 2 groups gathering between Cross_2 and Cross_4*
Group B (56.6-97.4% of contribution)	**Group C (64.5-98.0% of contribution)**
Cross_2: CS122A CS125B CS127D CS128A CS129D CS130A CS130C CS131B CS132D CS133A CS133C CS134A CS134C CS134D CS135D CS136D CS137C CS137D CS138A CS139A CS140C CS141B CS142D CS143A CS143B CS143C CS144A CS144C CS144D CS145A CS145B CS146A CS146B CS146D CS148B CS150A CS150B CS150D CS151B CS152A CS152D CS153B CS154B CS155B CS156B CS157D CS158B CS158D CS159B CS161C CS162A CS162D CS163B CS163C CS164D CS166D CS167B CS167C CS168B CS169B CS169D CS176D CS180A Cross_3: CS185D CS188D CS189A CS189B CS192B CS192C CS195C CS195D CS196C CS197E CS198A CS198B CS201C CS202C CS204A CS205C CS205D CS207C CS208C CS214A CS225C CS225D CS226D CS231C CS233C	Cross_2: CS133B CS138B CS147C CS148D* CS153C CS158C* CS160C CS162B Cross_4: CS241A CS241B CS241C CS241D CS248A CS248C CS250B CS251C CS259C* CS259D CS266D CS268D CS271B CS271C CS271C2 CS272A CS272A2 CS273C CS273D CS274A CS274C2 CS274D CS275A CS275D CS276A CS276B CS276C CS276D CS277A CS277B CS277B2 CS278A CS278B2 CS278D CS278D2 CS280A CS280B CS280C CS280C2 CS280D CS280D2 CS281B CS281D CS282C CS282C2 CS283D* CS284A CS284A2 CS284D CS284D2 CS286A CS286D CS287C CS289C* CS290A CS292D CS294C CS294D CS295A

REFERENCES

Adams, N.G., Trainer, V.L., Rocap, G., Herwig, R.P. and Hauser, L. 2009. Genetic population structure of *Pseudo-nitzschia pungens* (Bacillariophyceae) from the Pacific Northwest and the North Sea. *Journal of Phycology,* **45**: 1037-1045.

Akallal, R., Billard, C., Fresnel, J., Givernaud, T. and Mouradi, A. 2002. Phytoplankton of the Atlantic coasts of Morocco II. The genus *Pseudo-nitzschia* (Bacillariophyceae). *Cryptogamie Algologie,* **23**: 187-202.

Amato, A., Kooistra, W.H.C.F., Levialdi Ghiron, J.H., Mann, D.G., Pröschold, T. and Montresor, M. 2007. Reproductive isolation among sympatric cryptic species in marine diatoms. *Protist,* **158**: 193-207.

Amato, A., Lüdeking, A. and Kooistra, W.H.C.F. 2010. Intracellular domoic acid production in *Pseudo-nitzschia multistriata* isolated from the Gulf of Naples (Tyrrhenian Sea, Italy). *Toxicon,* **55**: 157-161.

Amos, W. 2009. Heterozygosity and mutation rate: evidence for an interaction and its implications. *BioEssays,* **32**: 82-90.

Arjan, G., J, Zeyl, C.W., Gerrish, P.J., Blanchard, J.L. and Lenski, R.E. 1999. Diminishing returns from mutation supply rate in asexual populations. *Science,* **283**: 404-406.

Astraldi, M. and Gasparini, G.P. 1994. The seasonal characteristics of the circulation in the Tyrrhenian Sea. In *Seasonal and Interannual Variability of the Western Mediterraneab Sea, Coastal Estuarine Studies* (P.E. La Violette, ed), pp. 115-134. Washington, D.C.: AGU.

Baldauf, S.L., Bhattacharya, D., Cockrill, J., Hugenholtz, P., Pawlowski, J. and Simpson, A.G.B. 2004. The Tree of Life. In *Assembling the Tree of Life* (J. Cracraft and M.J. Donoghue, eds). New York: Oxford University Press.

Barbera, M.A. and Petes, T.D. 2006. Selection and analysis of spontaneus reciprocal mitotic cross-over in *Sacharomyces cerevisiae. Proceeding of the National Academy of Sciences of the United States of America.,* **103**: 12819-12824.

Barka, S. 2007. Insoluble detoxification of trace metals in a marine copepod *Tigriopus brevicornis* (Müller) exposed to copper, zinc, nickel, cadmium, silver and mercury. *Ecotoxicology,* **16**: 491-502.

Bates, S.S. 2000. Domoic-acid-producing diatoms: another genus added! *Journal of Phycology,* **36**: 978-985.

Bates, S.S., Gaudet, J., Kaczmarska, I. and Ehrman, J.M. 2004. Interaction between bacteria and the domoic-acid-producing diatom *Pseudo-nitzschia multiseries* (Hasle) Hasle; can bacteria produce domoic acid autonomously? *Harmful Algae,* **3**: 11-20.

Beaumont, M.A. and Nichols, R.A. 1996. Evaluating loci for use in the genetic analysis of population structure. *Proceedings of the Royal Society B,* **363**: 1619-1626.

Bell, W.H. 1983. Bacterial utilization of algal extracellular products. 3. The specificity of algal-bacterial interaction. *Limnology and Oceanography,* **28**: 1131-1143.

Bracher, A., Vountas, M., Dinter, T., Burrows, J.P., Röttgers, R. and Peeken, I. 2009. Quantitative observation of cyanobacteria and diatoms from space using PhytoDOAS on SCIAMACHY data. *Biogeosciences,* **6**: 751-764.

Busse, L.B., Venrick, E.L., Antrobus, R., Miller, P.E., Vigilant, V., Silver, M.W., Mengelt, C., Mydlarz, L. and Prezelin, B.B. 2005. Domoic acid in phytoplankton and fish in San Diego, CA, USA. *Harmful Algae,* **5**: 91-101.

Callen, D.F., Thompson, A.D., Shen, Y., Philips, H.A., Richards, R.I., Mulley, J.C. and Sutherland, G.R. 1993. Incidence and Origin of "null" alleles in the (AC)n microsatellite markers. *American Journal of Human Genetics,* **52**: 922-927.

Carrada, G.C., Fresi, E., Marino, D., Modigh, M. and Ribera d'Alcalà, M. 1981. Structural analysis of winter phytoplankton in the Gulf of Naples. *Journal of Plankton Research,* **3**: 291-314.

Carvalho Pinto-Silva, C.R., Moukha, S., Matias, W.G. and Creppy, E.E. 2008. Domoic acid induces direct DNA damage and apoptosis in Caco-2 cells: Recent advances. *Environmental Toxicology,* **23**: 657-663.

Casteleyn, G., Adams, N.G., Vanormelingen, P., De Beer, A.-E., Sabbe, K. and Vyverman, W. 2009. Natural hybrids in the marine diatom *Pseudo-nitzschia pungens* (Bacillariophyceae): genetic and morphological evidence. *Protist,* **160**: 343-354.

Casteleyn, G., Chepurnov, V.A., Leliaert, F., Mann, D.G., Bates, S.S., Lundholm, N., Rhodes, L., Sabbe, K. and Vyverman, W. 2008. *Pseudo-nitzschia pungens* (Bacillariophyceae): A cosmopolitan diatom species? *Harmful Algae,* **7**: 241-257.

Cembella, A.D. 2003. Chemical ecology of eukaryotic microalgae in marine ecosystems. *Phycologia,* **42**: 420-447.

Chepurnov, V.A., Mann, D.G., Sabbe, K. and Vyverman, W. 2004. Experimental studies on sexual reproduction in diatoms. *International Review of Cytology,* **237**: 91-154.

Chepurnov, V.A., Mann, D.G., Von Dassow, P., Vanormelingen, P., Gillard, J., Inzé, D., Sabbe, K. and Vyverman, W. 2008. In search of new tractable diatoms for experimental biology. *BioEssays,* **30**: 692-702.

Cho, E.S., Hur, H.J., Byun, H.S., Lee, S.-G., Rhodes, L.L., Jeong, C.S. and Park, J.G. 2002. Monthly monitoring of domoic acid producer *Pseudo-nitzschia multiseries* (Hasle) Hasle using species-specific DNA probes and WGA lectins

and abundance of *Pseudo-nitzschia* species (Bacillariophyceae) from Chinhae Bay, Korea. *Botanica Marina,* **45**: 364-372.

Churro, C.I., Carreira, C.C., Rodrigues, F.J., Craveiro, S.C., Calado, A.J., Casteleyn, G. and Lundholm, N. 2009. Diversity and abundance of potentially toxic *Pseudo-nitzschia* Peragallo in Aveiro coastal lagoon, Portugal and description of a new variety, *P. pungens* var. *aveirensis* var. nov. *Diatom Research,* **24**: 35-62.

Coesel, S., Obornik, M., Varela, J., Falciatore, A. and Bowler, C. 2008. Evolutionary origins and functions of the carotenoid biosynthetic pathway in marine diatoms. *Public Library of Science ONE,* **3**: e2896.

Colegrave, N. 2002. Sex releases the speed limit on evolution. *Nature,* **420**: 664-666.

Collado-vides, L. 2002. Clonal architecture in marine macroalgae: ecological and evolutionary perspectives. *Evolutionary ecology,* **15**: 531-545.

Cornish-Bowden, A. 1985. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. *Nucleic Acids Research,* **13**: 3021-3030.

Costa, P.R., Rosa, R., Duarte-Silva, A., Brotas, V. and Sampayo, M.A.M. 2005. Accumulation, transformation and tissue distribution of domoic acid, the amnesic shellfish poisoning toxin, in the common cuttlefish, *Sepia officinalis. Aquatic Toxicology,* **74**: 82-91.

Creach, V., Ernst, A., Sabbe, K., Vanelslander, B., Vyverman, W. and Stal, L.J. 2006. Using quantitative PCR to determine the distribution of a semicryptic benthic diatom, *Navicula phyllepta* (Bacillariophyceae). *Journal of Phycology,* **42**: 1142-1154.

Croft, M.T., Warren, M.J. and Smith, A.G. 2006. Algae need their vitamins. *Eukaryotic Cell,* **5**: 1175-1183.

D'Alelio, D., Ribera D'Alcalà, M., Dubroca, L., Sarno, D., Zingone, A. and Montresor, M. 2010. The time for sex: A biennial life cycle in a marine planktonic diatom. *Limnology and Oceanography,* **55**: 106-114.

D'Alelio, D., Amato, A., Kooistra, W.H.C.F., Procaccini, G., Casotti, R. and Montresor, M. 2009a. Internal Transcribed Spacer polymorphism in *Pseudo-nitzschia multistriata* (Bacillariophyceae) in the Gulf of Naples: recent divergence or intraspecific hybridization? *Protist,* **160**: 9-20.

D'Alelio, D., Amato, A., Luedeking, A. and Montresor, M. 2009b. Sexual and vegetative phases in the planktonic diatom *Pseudo-nitzschia multistriata. Harmful Algae,* **8**: 225-232.

Devlin, B. and Risch, N. 1995. A comparison of linkage disequilibrium measures for fine-scale mapping. *Genomics,* **29**: 311-322.

Drouaud, J., Camilleri, C., Bourguignon, P.Y., Canaguier, A., Bérard, A., Vezon, D., Giancola, S., Brunel, D., Colot, V., Prum, B., Quesneville, H. and Mézard, C. 2006. Variation in crossing-over rates across chromosome 4 of *Arabidopsis thaliana* reveals the presence of meiotic recombination "hot spots". *Genome Research,* **16**: 106-114.

Ellengren, H. 2000. Heterogeneous mutation processes in human microsatellite DNA sequences *Nature Genetics,* **24**: 400-402.

Ellengren, H. 2004. Microsatellites: simple sequences with complex evolution. *Nature Reviews Genetics,* **5**: 435-443.

Estoup, A. and Angers, B. 1998. Microsatellites and minisatellites for molecular ecology: theoretical and empirical considerations. In *Advances in Molecular Ecology* (G.R. Carvalho, ed), pp. 55-86. Erice: IOS Press.

Estoup, A., Jarne, P. and Cornuet, J.M. 2002. Homoplasy and mutation model at microsatellite loci and their consequences for population genetic analysis. *Molecular Ecology,* **11**: 1591-1604.

Evans, K.M., Bates, S.S., Medlin, L.K. and Hayes, P.K. 2004. Microsatellite marker development and genetic variation in the toxic diatom *Pseudo-nitzschia multiseries* (Bacillariophyceae). *Journal of Phycology,* **40**: 911-920.

Evans, K.M., Chepurnov, V.A., Sluiman, H.J., Thomas, S.J., Spears, B.M. and Mann, D.G. 2009. Highly differential populations of the freshwater diatom *Sellaphora capitata* suggest limited dispersal and opportunities for allopatric speciation. *Protist,* **160**: 386-396.

Evans, K.M. and Hayes, P.K. 2004. Microsatellite markers for the cosmopolitan marine diatom *Pseudo-nitzschia pungens. Molecular Ecology Notes,* **4**: 125-126.

Evans, K.M., Kühn, S.F. and Hayes, P.H. 2005. High levels of genetic diversity and low levels of genetic differentiation in North Sea *Pseudo-nitzschia pungens* (Bacillariophyceae) populations. *Journal of Phycology,* **41**: 506-514.

Evans, K.M., Wortley, A.H. and Mann, D.G. 2007. An assessment of potential diatom "Barcode" genes (cox1,rbcL,18S and ITS rDNA) and their effectiveness in determining relationships in *Sellaphora* (Bacillariophyta). *Protist,* **158**: 349-364.

Féral, J.-P. 2002. How useful are the genetic markers in attempts to understand and manage marine biodiversity? *Journal of Experimental Marine Biology and Ecology,* **268**: 121-145.

Gagneux, P., Boesch, C. and Woodruff, D.S. 1997. Microsatellite scoring errors associated with noninvasive genotyping based on nuclear DNA amplified from shed hair. *Molecular Ecology,* **6**: 861-868.

Goss, R. and Jacob, T. 2010. Regulation and function of xanthophyll cycle-dependent photoprotection in algae. *Photosynthesis Research.*

Grieco, L., Tremblay, L.-B. and Zambianchi, E. 2005. A hybrid approach to transport processes in the Gulf of Naples: an application to phytoplankton and zooplankton population dynamics. *Continental Shelf Research,* **25**: 711-728.

Griffiths, A.J.F., Miller, J.H., Suzuki, D.T., Lewontin, R.C. and Gelbart, W.M. 1993. *An Introduction to Genetic Analysis.*, Vol. Fifth Edition. New York: Freeman W.H. and Compagny.

Grimsley, N., Péquin, B., Bachy, C., Moreau, H. and Piganeau, G. 2010. Cryptic sex in the smallest eukaryotic marine green alga. *Molecular Biology and Evolution,* **27**: 47–54.

Guillard, R.R.L. 1973. Division rates. In *Handbook of Phycological Methods: Culture Methods and Growth Treatment* (J.R. Stein, ed), p. 289. Cambridge, London: University Press.

Guillard, R.R.L. 1975. Culture of phytoplankton for feeding marine invertebrates. In *Culture of Marine Invertebrate Animals* (W.L. Smith and M.H. Chanley, eds), pp. 26-60: Plenum Press, New York, USA.

Guillard, R.R.L. and Ryther, J.H. 1962. Studies of marine planktonic diatoms. I. *Cyclotella nana* Hustedt and *Detonula confervacea* (Cleve). *Canadian Journal of Microbiology,* **8**: 229-239.

Hall, C.A.S., Stanford, J.A. and Hauer, E.R. 1992. The distribution and abundance of organisms as a consequence of energy balances along multiple environmental gradients. *Oikos,* **65**: 377-390.

Hall, T.A. 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. *Nucleic Acids Symposium Series,* **41**: 95-98.

Hancock, J.M. 1999. Microsatellites and other simple sequences: genomic context and mutational mechanisms. In *Microsatellites Evolution and Applications* (D.B. Goldstein and C. Schlötterer, eds). New York: Oxford University Press.

Hartl, D.L. and Clark, A.G. 1989. Population subdivision and migration. In *Principles of Population Genetics* (D.L. Hartl and A.G. Clark, eds). Subderland, Massachusetts: Sinauer Associates, Inc.

Hasle, G.R. 1994. *Pseudo-Nitzschia* as a genus distinct from *Nitzschia* (Bacillariophyceae). *Journal of Phycology,* **30**.

Hasle, G.R. 2002. Are most of the domoic acid-producing species of the diatom genus *Pseudo-nitzschia* cosmopolites? *Harmful Algae,* **1**: 137-146.

Henderson, S.T. and Petes, T.D. 1992. Instability of simple sequence DNA in *Saccharomyces cerevisiae. Molecular and Cellular Biology,* **12**: 2749-2757.

Hildebrand, M. 2008. Diatoms, Biomineralization processes, and Genomics. *Chemistry Review,* **108**: 4855-4874.

Hulot, F.D. and Huisman, J. 2004. Allelopathic interactions between phytoplankton species: the roles of heterotrophic bacteria and mixing intensity. *Limnology and Oceanography,* **49**: 1424-1434.

Kaczmarska, I., Ehrman, J.M., Bates, S.S., Green, D.H., Leger, C. and Harris, J. 2005. Diversity and distribution of epibiotic bacteria on *Pseudo-nitzschia multiseries* (Bacillariophyceae) in culture, and comparison with those on diatoms in native seawater. *Harmful Algae,* **4**: 725-741.

Kooistra, W.H.C.F. 2007. The evolution of the diatoms. In *Handbook of Biomineralization.* (E. Bäuerlein, ed). Weinheim: WILEY-VCH Verlag GmbH & Co.

Kooistra, W.H.C.F., Boele-Bos, S.A. and Stam, W.T. 1991. A method for obtaining axenic algal cultures using antibiotic cefotaxime with emphasis on *Cladophoropsis membranacea* (Chlorophyta). *Journal of Phycology,* **27**: 656-658.

Kooistra, W.H.C.F., Forlani, G. and De Stefano, M. 2009. Adaptations of araphid pennate diatoms to a planktonic existence. *Marine Ecology,* **30**: 1-15.

Kooistra, W.H.C.F., Gersonde, R., Medlin, L.K. and Mann, D.G. 2007. The origin and evolution of the diatoms: their adaptation to a planktonic existence. In *Evolution of primary producers in the sea* (P.G. Falkowski and A.H. Knoll, eds): Elsevier.

Kooistra, W.H.C.F., Sarno, D., Balzano, S., Gu, H., Andersen, R.A. and Zingone, A. 2008. Global diversity and biogeography of *Skeletonema* species (Bacillariophyta). *Protist,* **159**: 177-193.

Kröger, N. and Poulsen, N. 2008. Diatoms - From cell wall biogenesis to nanotechnology. *Annual Review Genetic,* **42**: 83-107.

Kroth, P.G., Chiovitti, A., Gruber, A., Martin-Jezequel, V., Mock, T., Parker, M.S., Stanley, M.S., Kaplan, A., Caron, L., Weber, T., Maheswari, U., Armbrust, E.V. and Bowler, C. 2008. A model for carbohydrate metabolism in the diatom *Phaeodactylum tricornutum* deduced from comparative whole genome analysis. *Public Library of Science ONE,* **3**: e1426.

Lafontaine, D.L.J. and Tollervey, D. 2001. The function and synthesis of ribosomes. *Nature,* **2**: 514-520.

Lavaud, J. 2007. Fast regulation of photosynthesis in diatoms: Mechanisms, Evolution and Ecophysiology. *Functional Plant Science and Biotechnology,* **1**: 267-287.

Lee, R.E. 1999. *Phycology.* Cambridge, New York, Melbourne, Madrid: Cambridge University Press.

Leppard, G.G. 1995. The characterization of algal and microbial mucilages and their aggregates in aquatic ecosystems. *Science of the Total Environment,* **165**: 103-131.

Levinson, G. and Gutman, G.A. 1987. Slipped-strand mispairing: A major mechanism for DNA sequence evolution. *Molecular Biology and Evolution,* **4**: 203-221.

Lewin, J.C., Lewin, R.A. and Philpott, D.E. 1958. Observations on *Phaeodactylum tricornutum*. *Journal of General Microbiology,* **18**: 418-426.

Litchman, E., De Tezanos Pinto, P., Klausmeier, C.A., Thomas, M.K. and Yoshiyama, K. 2010. Linking traits to species diversity and community structure in phytoplankton. *Hydrobiologia,* **653**: 15-28.

Lynch, M., Sung, W., Morris, K., Coffey, N., Landry, C.R., Dopman, E.B., Dickinson, W.J., Okamoto, K., JKulkarni, S., Hartl, D.L. and Thomas, W.K. 2008. A genome-wide view of the spectrum of spontaneous mutations in yeast. *Proceeding of the National Academy of Sciences of the United States of America,* **105**: 9272-9277.

MacDonald, J.D. 1869. On the structure of the diatomaceous frustule, and its genetic cycle. *The Annals and Magazine of Natural History,* **Ser. 4, 3**: 1-8.

MacIntyre, H.L., Kana, T.M. and Geider, R.J. 2000. The effect of water motion on short-term rates of photosynthesis by marine phytoplankton. *Trends in Plant Science,* **5**: 12-17.

Mann, D.G. 1994. Auxospore formation, reproductive plasticity and cell structure in *Navicula ulvacea* and resurrection of the genus *Dickieia* (Bacillariophyta). *European Journal of Phycology,* **29**: 141-157.

Mann, D.G., Chepurnov, V.A. and Idei, M. 2003. Mating system, sexual reproduction, and auxosporulation in the anomalous raphid diatom *Eunotia* (Bacillariophyta). *Journal of Phycology,* **39**: 1067-1084.

Meldrum, B.S. 2000. Glutamate as a Neurotransmitter in the Brain: Review of Physiology and Pathology. *American Society for Nutritional Sciences.*, **130**: 1007-1015.

Menna, M., Mercatini, A., Uttieri, M., Buonocore, B. and Zambianchi, E. 2007. Wintertime transport processes in the Gulf of Naples investigated by HF radar measurements of surface currents. *Il Nuovo Cimento,* **30**: 605-622.

Metzgar, D. and Wills, C. 2000. Evidence for the Adaptative Evolution of Mutation Rates. *Cell,* **101**: 581-584.

Miller, P.E. and Scholin, C.A. 1998. Identification and enumeration of cultured and wild *Pseudo-nitzschia* (Bacillariophyceae) using species-specific LSU rRNA-targeted fluorescent probes and filtered-based whole cell hybridization. *Journal of Phycology,* **34**: 371-382.

Miller, P.E. and Scholin, C.A. 2000. On detection of *Pseudo-nitzschia* (Bacillariophyceae) species using whole cell hybridization: sample fixation and stability. *Journal of Phycology,* **36**: 238-250.

Miralles, R., Gerrish, P.J., Moya, A. and Elena, S.F. 1999. Clonal interference and the evolution of RNA viruses. *Science,* **285**: 1745-1747.

Montresor, M., Sgrosso, S., Procaccini, G. and Kooistra, W.H.C.F. 2003. Intraspecific diversity in *Scrippsiella trochoidea* (Dinophyceae): evidence for cryptic species. *Phycologia,* **42**: 56-70.

Moschandreou, K.K. and Nikolaidis, G. 2010. The genus *Pseudo-nitzschia* (Bacillariophyceae) in Greek coastal waters. *Botanica Marina,* **53**: 159-172.

Nezan, E. and Chomerat, N. 2007. Identification of *Pseudo-nitzschia multistriata* and *P. subpacifica* from French waters. Were they part of the cryptic flora. *Harmful Algae News,* **35**: 5-6.

Obernosterer, I. and Herndl, G.J. 1995. Phytoplankton extracellular release and bacterial growth: dependence on the inorganic N:P ratio. *Marine Ecology Progress Series,* **116**: 247-257.

Oetjen, K. and Reusch, T.B.H. 2007. Genome scans detect consistent divergent selection among subtidal vs. intertidal populations of the marine angiosperm *Zostera marina. Molecular Ecology,* **16**: 5156-5167.

Orive, E., Laza-Martinez, A., Seoane, S., Alonso, A., Andrade, R. and Miguel, I. 2010. Diversity of *Pseudo-nitzschia* in the southeastern Bay of Biscay. *Diatom Research,* **25**: 125-145.

Orlova, T.Y., Stonik, I.V., Aizdaicher, N.A., Bates, S.S., Léger, C. and Fehling, J. 2008. Toxicity, morphology and distribution of *Pseudo-nitzschia calliantha, P. multistriata* and *P. multiseries* (Bacillariophyta) from the northwestern Sea of Japan. *Botanica Marina,* **51**: 297-306.

Orsini, L., Sarno, D., Procaccini, G., Poletti, R., Dahlmann, J. and Montresor, M. 2002. Toxic *Pseudo-nitzschia multistriata* (Bacillariophyceae) from the Gulf of Naples: morphology, toxin analysis and phylogenetic relationships with other *Pseudo-nitzschia* species. *European Journal of Phycology,* **37**: 247-257.

Ouborg, N.J., Piquot, Y. and Van Groenendael, J.M. 1999. Population genetics, molecular markers and the study of dispersal in plants. *Journal of Ecology,* **87**: 551-568.

Pal, C., Papp, B. and Hurst, L.D. 2001. Does the recombination rate affect the efficiency of purifying selection? The yeast genome provides a partial answer. *Molecular Biology and Evolution,* **18**: 2323-2326.

Peakall, R. and Smouse, P.E. 2006. GENEALEX 6: genetic analysis in Excel. Population genetic software for teaching and research. *Molecular Ecology Notes,* **6**: 288-295.

Pfitzer, E. 1869. Über den Bau und Zellteilung der Diatomeen . *Botanische Zeitung,* **27**: 774-776.

Pfrender, M.E. and Lynch, M. 2000. Quantitative genetic variation in *Daphnia*: Temporal changes in genetic architecture. *Evolution,* **54**: 1502-1509.

Pritchard, J.K., Stephens, M. and Donnelly, P. 2000. Inference of population structure using multilocus genotype data. *Genetics,* **155**: 945-959.

Procaccini, G. and Maltagliati, F. 2004. Methodological approaches to the analysis of genetic diversity in benthic organisms. (trans.) B.M.M. II. In *Methods for sampling and study of Mediterranean marine benthos.* (Gambi M.C. and M. Dappiano, eds), pp. 481-519.

Pugnetti, A., Camatti, E., Mangoni, O., Morabito, G., Oggioni, A. and Saggiomo, V. 2006. Phytoplankton production in Italian freshwater and marine ecosystems: State of the art and perspectives. *Chemistry and Ecology,* **22**: S49-S69.

Quijano-Scheggia, S., Garcés, E., Andree, K., Fortuño, J.M. and Camp, J. 2009. Homothallic auxosporulation in *Pseudo-nitzschia brasiliana* (Bacillariophyta). *Journal of Phycology,* **45**: 100-107.

Quijano-Scheggia, S., Garcés, E., Flo, E., Fernandez-Tejedor, M., Diogène, J. and Camp, J. 2008a. Bloom dynamics of the genus *Pseudo-nitzschia* (Bacillaryophyceae) in two coastal bays (NW Mediterranean Sea). *Scientia Marina,* **72**: 577-590.

Quijano-Scheggia, S., Garces, E., Sampedro, N., van Lenning, K., Flo, E., Andree, K., Fortuno, J.M. and Camp, J. 2008b. Identification and characterisation of the dominant *Pseudo-nitzschia* species (Bacillariophyceae) along the NE Spanish coast (Catalonia, NW Mediterranean). *Scientia Marina,* **72**: 343-359.

Quijano-Scheggia, S., Garces, K., Van Lenning, K., Sampedro, N. and Camp, J. 2006. First detection of diatom *Pseudo-nitzschia brasiliana* (non toxic) and its relative *P. multistriata* (presumably toxic) in the NW Mediterranean Sea. *Harmful Algae News,* **31**: 5.

Reich, D.E. and Goldstein, D.B. 1999. Estimating the age of mutations using variation at linked markers. In *Microsatellites - Evolution and Applications* (D.B. Goldstein and C. Schlötterer, eds). New York: Oxford University Press.

Rhodes, L.L., Adamson, J. and Scholin, C. 2000. *Pseudo-nitzschia multistriata* (Bacillariophyceae) in New Zealand. *New Zealand Journal of Marine and Freshwater Research,* **34**: 463-467.

Rhodes, L.L., Scholin, C., Garthwaite, I., Haywood, A. and Thomas, A. 1998. Domoic acid producing *Pseudo-nitzschia species* educed by whole cell DNA probe-based and immunochemical assays. Paper presented at the Harmful Algae, Vigo, 1998.

Ribera d'Alcalà, M., Conversano, F., Corato, F., Licandro, P., Mangoni, O., Marino, D., Mazzocchi, G., Modigh, M., M., M., Nardella, M., Saggiomo, V., Sarno, D. and Zingone, A. 2004. Seasonal patterns in plankton communities in a pluriannual time series at a coastal Mediterranean site (Gulf of Naples): an attempt to discern recurrences and trends. *Scientia Marina,* **68**: 65-83.

Rogall, E. 1939. Über den Feinbau der Kieselmembran der Diatomeen. *Planta,* **29**: 279-291.

Round, F.E., Crawford, R.M. and Mann, D.G. 1990. *The diatoms. Biology and morphology of the genera*. Cambridge: Cambridge University Press.

Rozen, S. and Skalestsky, H. 2000. Primer3 on the WWW for general users and for biologist programmers. *Methods in Molecular Biology,* **132**: 365-386.

Rue, E. and Bruland, K. 2001. Domoic acid binds iron and copper: a possible role for the toxin produced by the marine diatoms *Pseudo-nitzschia. Marine Chemistry,* **76**: 127-134.

Rynearson, T.A. and Armbrust, E.V. 2000. DNA fingerprinting reveals extensive genetic diversity in a field population of the centric diatom *Ditylum brightwellii. Limnology and Oceanography,* **45**: 1329-1340.

Rynearson, T.A. and Armbrust, E.V. 2005. Maintenance of clonal diversity during a spring bloom of the centric diatom *Ditylum brightwellii. Molecular Ecology,* **14**: 1631-1640.

Rynearson, T.A. and Armbrust, V.E. 2004. Genetic differentiation among populations of the planktonic marine diatom *Ditylum brightwellii* (Bacillariophyceae). *Journal of Phycology,* **40**: 34-43.

Rynearson, T.A., Lin, E.O. and Armbrust, E.V. 2009. Metapopulation structure in the planktonic diatom *Ditylum brightwellii* (Bacillariophyceae). *Protist,* **160**: 111-121.

Sarno, D. and Dahlman, J. 2000. Production of domoic acid in another species of *Pseudo-nitzschia*: *P. multistriata* in the Gulf of Naples (Mediterranean Sea). *Harmful Algae News,* **21**: 5.

Sarno, D., Kooistra, W.H.C.F., Medlin, L.K., Percopo, I. and Zingone, A. 2005. Diversity in the genus *Skeletonema* (Bacillariophyceae). II. An assessment of the taxonomy of *S. costatum-like species* with the description of four new species. *Journal of Phycology,* **41**: 151-176.

Scholin, C.A., Marin III, R., Miller, P.E., Doucette, G.J., Powell, C.L., Haydock, P., Howard, J. and Ray, J. 1999. DNA probes and a receptor-binding assay for detection of *Pseudo-nitzschia* (Bacillariophyceae) species and domoic acid activity in cultured and natural samples. *Journal of Phycology,* **35**: 1356-1367.

Schubert, H. and Forster, R.M. 1997. Sources of variability in the factors used for modelling primary productivity in eutrophic waters. *Hydrobiologia,* **349**: 75-85.

Selkoe, K.A. and Toonen, R.J. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters,* **9**.

Sell, A.F. and Overback, J. 1992. Exudates: phytoplankton-bacterioplankton interactions in Plusssee. *Journal of Plankton Research,* **14**: 1199-1215.

Sgrosso, S., Esposito, F. and Montresor, M. 2001. Temperature and daylength regulate encystment in calcareous cyst-forming dinoflagellates. *Marine Ecology Progress Series,* **211**: 77-87.

Shinde, D., Lai, Y., Sun, F. and Arnheim, N. 2003. *Taq* DNA polymerase slippage mutation rates measured by PCR and quasi-likelihood analysis: (CA/GT)*n* and (A/T)*n* microsatellites. *Nucleic Acids Research,* **31**: 974-980.

Shipley, P. 2003. Micro-Checker, Version 2.2.3. University of Hull, Hull, UK, URL http://www.microchecker.hull.ac.uk/.

Siano, R. 2008. *The phytoplankton of the Campania coasts: an ecological and taxonomic study*: PhD thesis. University of Messina, Italy.

Sniegowski, P.D., Gerrish, P.J., Johnson, T. and Shaver, A. 2000. The evolution of mutation rates: separating causes from consequences. *BioEssays,* **22**: 1057-1066.

Stacey, K.A. 1994. Recombination. In *The Encyclopedia of Molecular Biology* (J. Kendrew and E. Lawrence, eds), pp. 945-950. Oxford: Blackwell Science.

Storlazzi, A., Xu, L., Cao, L. and Kleckner, N. 1995. Crossover and noncrossover recombination during meiosis: Timing and pathway relationships. *Proceeding of the National Academy of Sciences of the United States of America,* **92**: 8512-8516.

Strand, M., Prolla, T.A., Liskay, R.M. and Petes, T.D. 1993. Destabilization of tracts of simple repetitive DNA in yeast by mutations affecting DNA mismatch repair. *Nature,* **365**: 274-276.

Suzuki, L. and Johnson, C.H. 2001. Algae know the time of day: circadian and photoperiodic programs. *Journal of Phycology,* **37**: 933-942.

Takano, H. 1993. Marine diatom *Nitzschia multistriata* sp. nov. common at inlets of southern Japan. *Diatom,* **8**: 39-41.

Takano, H. 1995. *Pseudo-nitzschia multistriata* (Takano) Takano, a new combination for the pennate diatom *Nitzschia multistriata* Takano. *Diatom,* **10**: 73-74.

Tillmann, U., Alpermann, T., John, U. and Cembella, A.D. 2008. Allelochemical interactions and short-term effects of the dinoflagellate *Alexandrium* on selected photoautotrophic and hetetrophic protists. *Harmful Algae,* **7**: 52-64.

Tillmann, U., John, U. and Cembella, A.D. 2007. On the allelochemical potency of the marine dinoflagellate *Alexandrium ostenfeldii* against heterotrophic and autotrophic protists. *Journal of Plankton Research,* **29**: 527-543.

Uribe, P. and Espejo, R.T. 2003. Effect of associated bacteria on the growth and toxicity of *Alexandrium catenella*. *Environmental Microbiology,* **69**: 659-662.

Valière, N. 2002. GIMLET: a computer program for analysing genetic individual identification data. *Molecular Ecology Notes,* **2**: 377-379.

Vigilant, V. and Silver, M.W. 2007. Domoic acid in benthic flatfish on the continental shelf on Monterey Bay, California, USA. *Marine Biology,* **151**: 2053-2062.

Vigouroux, Y., Jaqueth, J.S., Matsuoka, Y., Smith, O.S., Beavis, W.D., Smith, S.C. and Doebley, J. 2002. Rate and pattern of mutation at microsatellite loci in Maize *Molecular Biology and Evolution,* **19**: 1251-1260.

von Dassow, P., Chepurnov, V.A. and Armbrust, E.V. 2006. Relationships between growth rate, cell size, and induction of spermatogenesis in the centric diatom *Thalassiosira weissflogii* (Bacillariophyta). *Journal of Phycology,* **42**: 887-899.

Wattier, R., Dallas, J.F., Destombe, C., Saumitou-Laprade, P. and Valero, M. 1997. Single locus microsatellites in *Gracilariales* (Rhodophyta): High level of genetic variability with *Gracilaria gracilis* and conservation in related species. *Journal of Phycology,* **33**: 868-880.

Wattier, R., Engel, C.R., Saumitou-Laprade, P. and Valero, M. 1998. Short allele dominance as a source of heterozygote deficiency at microsatellite loci: experimental evidence at the dinucleotide locus Gv1CT in *Gracilaria gracilis* (Rhodophyta). *Molecular Ecology,* **7**: 1569-1573.

Wenzl, S., Hett, R., Richthammer, P. and Sumper, M. 2008. Silacidins: Highly acidic phosphopeptides from diatom shells assist in silica precipitation in vitro. *Angewandte Chemie International Edition,* **47**: 1729-1732.

White, T.J., Bruns, T., Lee, S. and Taylor, J. 1990. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In *PCR protocols* (M.A. Innis, D.H. Gelfand, J.J. Sninsky and T.J. White, eds), pp. 315-322. New York: Academic Press.

Wierdl, M., Dominska, M. and Petes, T.D. 1997. Microsatellite instability in Yeast: dependence on the length of the microsatellite. *Genetics,* **146**: 769-779.

Zhu, F., Massana, R., Not, F., Marie, D. and Vaulot, D. 2005. Mapping of picoeucaryotes in marine ecosystems with quantitative PCR of the 18S rRNA gene. *Fems Microbiology Ecology,* **52**: 79-92.

Zingone, A., Dubroca, L., Iudicone, D., Margiotta, F., Corato, F., Ribera d'Alcalà, M., Saggiomo, V. and Sarno, D. 2010. Coastal phytoplankton do not rest in winter. *Estuaries and Coasts,* **33**: 342-361.

Zingone, A. and Enevoldsen, H.O. 2000. The diversity of harmful algal blooms: a challenge for science and management. *Ocean and Coastal Management,* **43**: 725-748.

Zingone, A., Siano, R., D'Alelio, D. and Sarno, D. 2006. Potentially toxic and harmful microalgae from coastal waters of the Campania region (Tyrrhenian Sea, Mediterranean Sea). *Harmful Algae,* **5**: 321-337.